Three-Dimensional Reconstruction of Fire from Images

by

Samuel William Hasinoff

A thesis submitted in conformity with the requirements
for the degree of Master of Science
Graduate Department of Computer Science
University of Toronto

Abstract

Three-Dimensional Reconstruction of Fire from Images

Samuel William Hasinoff
Master of Science
Graduate Department of Computer Science
University of Toronto
2002

Interesting visual simulations have been proposed for physical phenomena such as fire and smoke, but scant attention has been paid to reconstructing models of these phenomena from multiple images of real scenes. Here we study the problem of reconstructing three-dimensional models of fire from synchronized images or video sequences. Using a physically-based imaging model, we reduce the reconstruction of fire to the little-studied problem of computerized tomography with an extremely limited number of views (as few as two). We analyse the space of photo-consistent reconstructions, propose a novel reconstruction method for fire, and develop an image-based method for rendering the reconstructions. The reconstruction method involves parameterizing the fire with a linear superposition of Gaussians, estimating the density according to the input images, and applying a stochastic resampling scheme to refine the reconstruction. Experimental results are presented for a variety of synthetic and real fire datasets.

Acknowledgements

I'm profoundly indebted to Kyros Kutulakos for his thoughtful supervision, helpful suggestions, and many challenging discussions. Thanks as well to Allan Jepson for his valuable comments and assistance editing a late draft of this thesis. I would also like to acknowledge feedback from the members of the Dynamic Graphics Project, who saw an early version of this work as an informal talk, as well as interesting discussions with Sam Roweis, Chakra Chennubhotla, Karan Singh, and Radford Neal. The student regulars of the vision lab (John, Jennifer, Gustavo, Diego, Jing-Rui, and Chinh) deserve special mention for their collegial support and creating a stimulating work environment. Special thanks to Maxim Trokhimtchouk for his help formalizing the analysis of the solution space, and Alex Levinshtein for his help setting up the cameras. The CAMPFIRE dataset was graciously provided by Duc Nguyen of Stanford University. This work was supported in part by NSERC PGSA-232235-2000.

Contents

List of Figures

Chapter 1

Introduction

> The fire is the main comfort of the camp, whether in summer or winter, and is about as ample at one season as at another. It is as well for cheerfulness as for warmth and dryness.
>
> *The Maine Woods* (1864)
> Henry David Thoreau

> Fire is good. Fire . . . fire is our friend.
>
> Young Frankenstein (1974)

With its terrifying power and beauty, fire has always fascinated. As early as 1.5 million years ago, primitive man began harnessing fire as a tool for survival [18]. Because fire has such a wide range of practical uses, this was undoubtedly one the most important innovations of the era. Fire can be used to provide light and warmth, cook food, and enable long-distance communication. According to the Greek legend of Prometheus, fire was originally a secret stolen from the gods. Even to this day, fire retains an element of this otherworldliness and mystery.

The problem engaged by this thesis is the 3D reconstruction of fire from synchronized images or video sequences. Standard techniques from computer vision and scanning devices like laser rangefinders do not apply to this problem because of the transparent and volumetric nature of fire, and the practically-limited number of input views.

1.1 Motivation

The problem of 3D reconstruction of static, opaque objects from images is well understood by the computer vision community, and great progress has recently been made in this direction [102, 64, 89]. However, much less is known about how to capture 3D

representations of more general phenomena, especially dynamic ones. Specifically, few reconstruction methods have addressed the reconstruction of objects with mirror-like surfaces, transparency, or refractive properties.

Current reconstruction methods are particularly ill-suited to handling natural phenomena such as fire, smoke, and other fluids. The challenge is that these phenomena are intrinsically transparent, volumetric, and dynamic. As such, very little is known about how these situations can be handled starting from just 2D images. Even the question of what representation ought to be used for the phenomena, in order to be both useful and recoverable from images, is not clear. In this work, we provide some answers along these lines for the phenomenon of fire.

From the perspective of computer graphics, there has been a great deal of work on how to model and simulate natural phenomena (such as fire) for rendering. Although large strides have been made in this area [82, 65], we believe that the ultimate realism will come from capturing and manipulating representations from the real world.

While previous approaches have directly reused images of natural phenomena to produce realistic animations [100], the ultimate objective ought to be capturing 3D representations that are flexible enough to be manipulated. In this respect, capturing a dynamic 3D model of fire, representing a new volume every sixtieth of a second, would open up a whole set of new possibilities:

- Reusing the 3D captured data to create new simulations and animations.
- Training physical simulators with the reconstruction results so that they more accurately replicate the observed physical processes.
- Embedding the 3D captured fire into new virtual environments, or compositing the fire into existing video footage.
- Using the 3D captured data in conjunction with the input images to introduce principles from image-based rendering to the modelling and simulation of natural phenomena.

Fire is one specific context where these principles can be applied. We believe that image-based methods will eventually prove useful for modelling more general natural phenomena as well.

1.2 Summary of Contributions

- We introduce the problem of fire reconstruction to the computer vision community, with a review of related previous work in the combustion, computer graphics, and computer vision literature.

- We formalize the problem of image-based fire reconstruction from first principles, using elements from the optics of fire. We show that under a physically-based imaging model of fire, the problem is equivalent to computerized tomography. However, practical limits on the number of synchronized cameras available means that many fewer input images are available than in typical tomography applications.

- We present a theoretical analysis of the inherent ambiguity of the reconstruction problem in terms of fundamental switching operations. For the two-camera reconstruction problem, we also devise two analytic solutions representing extremes of the ambiguous solution space. This analysis is relevant not just for fire, but for tomography in general.

- We propose a novel method for the image-based reconstruction of fire. This method is based on parameterizing the fire with blobs, applying optimization techniques, and using stochastic resampling to refine the initial reconstruction.

- We also suggest a new rendering algorithm for fire that combines aspects of both volumetric and image-based rendering. This approach is also applicable to more general volumetric datasets.

1.3 Structure of the Thesis

The rest of this thesis is organized as follows. In Chapter 2, the relevant physics of fire is described, along with basic fire measurement techniques, and different image formation models for fire. Chapter 3 presents fire from the perspective of computer graphics. This includes methods for the modelling and simulation of fire (with visual if not physical realism), as well as a discussion of rendering techniques for fire. Then, in Chapter 4, fire is presented from the perspective of computer vision. After introducing computerized tomography, we describe the perception of transparency and review various computer vision techniques that account for transparency. Finally, in Chapter 5 we present our

analysis of the space of solutions that project to the input images, then give the details of our new reconstruction method, as well our algorithm for the image-based rendering of fire. Chapter 6 is devoted to general conclusions.

Chapter 2

Physics of Fire

> We must pass like smoke or live within the spirit's fire;
> For we can no more than smoke unto the flame return
> If our thought has changed to dream, our will unto desire,
> As smoke we vanish though the fire may burn.
>
> "Immortality" (l. 1–4)
> George William Russell

> Foul water will quench fire.
>
> Sixteenth-century British proverb

With the advent of modern chemistry, fire became understood as a combustion reaction requiring the participation of oxygen [120]. Combustion reactions always release heat, because as reactants are decomposed into simpler products some of the energy stored in chemical bonds is freed. Due to the high temperatures generated by combustion, these reactions occur mainly in the gaseous phase. Flame is defined intuitively as the hot, luminous gases associated with a combustion reaction.

2.1 Basic Physical Properties

2.1.1 Fire as Radiation

While fire always produces heat, fire is very often a source of light as well. When our eyes perceive flame, they are registering visible light radiation emitted from the hot gaseous and particulate reaction products. Fire can be characterized in terms of temperature and chemical composition based on its emission spectra [120].

The starting point for understanding radiation is the blackbody, an ideal object defined as the perfect emitter of radiation. In practice, blackbodies approximated by the

radiation escaping from a hole drilled in the side of a hot oven. Note that flame is well-approximated by blackbody radiation.

The Planck equation [31] describes how the spectral radiance L of an object at a given temperature T varies over a range of different wavelengths λ:

$$L(T, \lambda) = \frac{2\xi_\lambda \pi c^2}{\lambda^5 \left(e^{hc/\lambda k_B T} - 1\right)}. \tag{2.1}$$

This equation includes several constants related to radiation: the speed of light c, Planck's constant h, and Boltzmann's constant k_B. The spectral emissivity $0 \leq \xi_\lambda \leq 1$ is an empirical measure relating how closely the object approximates an ideal blackbody of the same temperature. When $\xi_\lambda = 1$, the object is emitting maximum radiance and behaving like a true blackbody. At sufficiently high temperatures all bodies become good radiators, explaining why objects placed in fire tend to appear closer in colour.

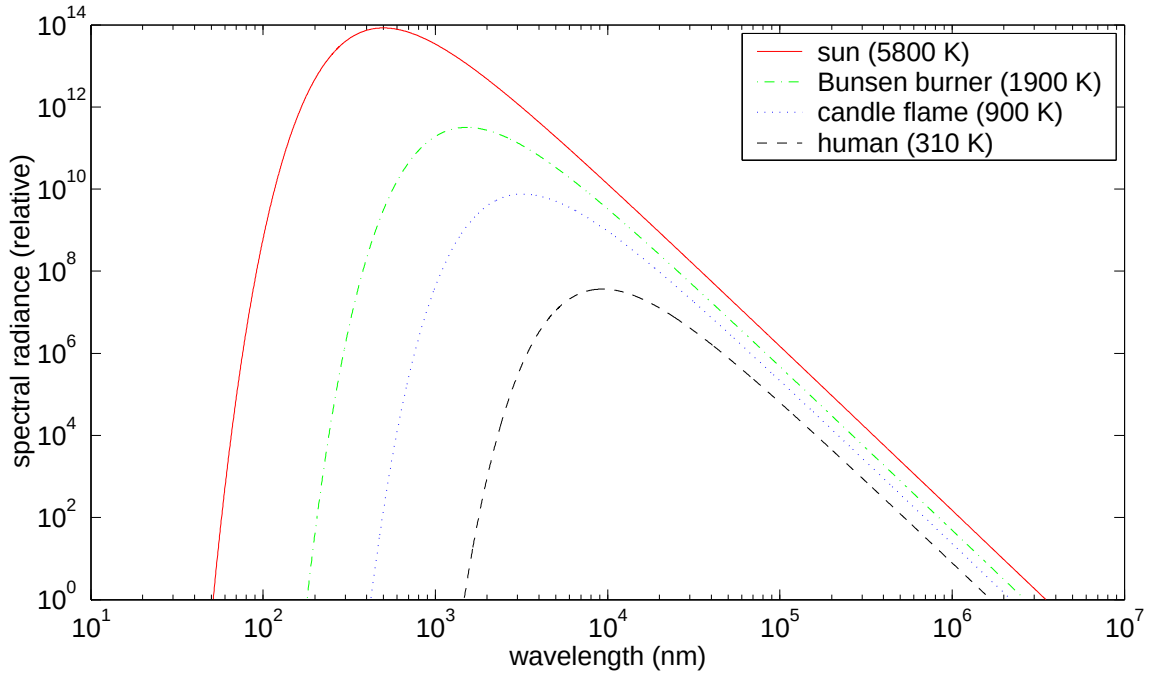


Figure 2.1: Wavelength distribution of blackbodies of various temperatures.

The spectral radiance for an ideal blackbody is strongly peaked, especially at high temperatures (Figure 2.1). It follows from Planck's equation (2.1) that the maximum spectral radiance for a blackbody will occur at wavelength

$$\lambda_{max} = \frac{2987 \text{ K} \cdot \mu\text{m}}{T}.$$

This relationship was first observed empirically as Wien's Law. Now integrating the Planck equation for a blackbody over all wavelengths gives the Stefan-Boltzmann equation:

$$L(T) = \int_0^\infty L(T, \lambda)\, \mathrm{d}\lambda = \left(\frac{2\pi^4 k_B^4}{15 c^2 h^3} \right) T^4.$$

This equation demonstrates that total radiance of a blackbody increases with the fourth power of temperature.

While blackbodies are perfect emitters of radiation by definition, they are also perfect absorbers of radiation. This is a consequence of Kirchoff's law, which states that absorptivity α_λ and emissivity ξ_λ are equal under local thermodynamic equilibrium. Absorptivity is defined analogously to emissivity as the fraction of absorbance relative to the absorbance of an ideal blackbody at the same temperature. This relationship will prove useful later for relating the absorptivity predicted by spectroscopy to the emissivity of luminous flame.

The fundamental equation of spectroscopy is the Beer-Lambert Law, which describes the reduction of flux density of a beam of light passing through some medium. The derivation presented here, as illustrated by Figure 2.2, involves integrating over infinitesimal unit area slabs using an approximate model for absorption.

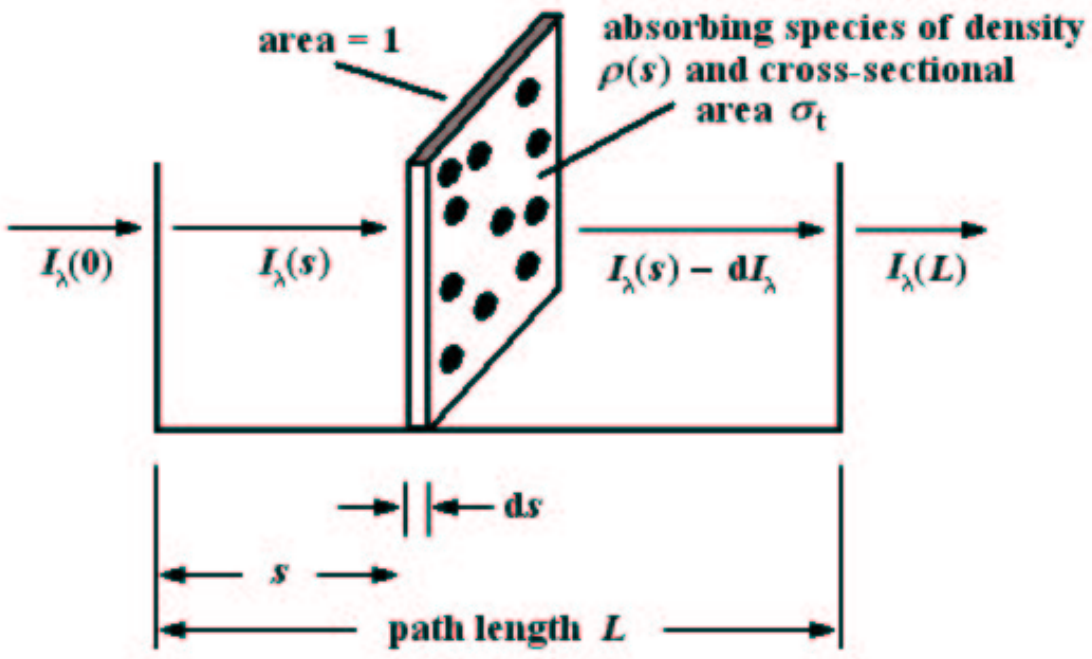


Figure 2.2: Derivation of the Beer-Lambert Law.

We approximate absorption by imagining that the infinitesimal slabs are blocked by regions of complete opaqueness which reduce the irradiance by an amount proportional to their area. The total area of these opaque regions can be expressed in terms of the density of the medium ρ and a constant of proportionality σ_t known as the extinction cross-section. The number of particles in an infinitesimal unit area slab is $\rho(s)\mathrm{d}s$, giving the total area of extinction as $\sigma_t\rho(s)\mathrm{d}s$. The differential falloff of the irradiance I_λ at position s is therefore:

$$\mathrm{d}I_\lambda = -\sigma_t\,\rho(s)\,I_\lambda(s)\,\mathrm{d}s.$$

Note that the extinction cross-section depends on the composition of the medium but is assumed not to vary with density. Now integrating over a path of length L yields

$$\ln I_\lambda(L) - \ln I_\lambda(0) = -\sigma_t \int_0^L \rho(s)\,\mathrm{d}s.$$

Stretching the definition of absorptivity somewhat, by assuming that a medium composed of blackbodies would absorb all of the incident light falling on the area of extinction, we obtain:

$$\begin{aligned} \alpha_\lambda &= \frac{I_\lambda(0) - I_\lambda(L)}{I_\lambda(0)} \\ &= 1 - \exp\left(-\sigma_t \int_0^L \rho(s)\,\mathrm{d}s\right). \end{aligned} \tag{2.2}$$

In fact, Equation (2.2) simplifies even further if we assume a constant density $\rho(s) = \rho$ throughout the medium. This gives the Beer-Lambert Law in its traditional form,

$$\alpha_\lambda = 1 - \exp(-\sigma_t\rho L). \tag{2.3}$$

The Beer-Lambert Law has a number of limitations. It tends to break down at very high and low concentrations, especially if the material is highly scattering. If the light is very intense, nonlinear optical processes can also cause variances. Strictly speaking, the Beer-Lambert Law should only be used for monochromatic radiation, but in practice it is often used for broadband radiation as well [36].

2.1.2 Refraction in Fire

The refraction of light in flames is generally small, with deflections less than 0.1° [101, 3]. Nonetheless, this amount of refraction is sometimes enough to produce visually significant distortions of the background, similar to the distortions caused by heat waves over hot

asphalt. This phenomenon is even more obvious when the fire is viewed over time and the turbulent nature of the refractive field is revealed through the chaotic warping of the background. In many experiments involving fire, the background is chosen to be uniformly black for convenience [3, 29, 31]. Furthermore, the mirage-like optical illusion that causes images of flame to be slightly displaced from their true positions is usually ignored, except when these deflection measurements are recorded explicitly in order to measure the refractive index field (§2.2.2).

The familiar Snell's Law [120] describes the instantaneous change in the direction of a ray of light at a sharp interface between two homogeneous media with different refractive indices n_1, n_2. The bending of light is described by the different angles θ_1, θ_2 made between the incoming and outgoing rays with the normal to the interface respectively:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2.$$

For spatial regions with a continuously varying refractive index the situation is more challenging. The light may follow an oddly curved path if the refractive index integrated over that path is a local minimum. This general case can be derived using Fermat's principle: the actual path between two points taken by a beam of light is the one which is traversed in the least time. More precisely, this path need only be stationary with respect to small variations, so that in special situations the path taken may even be a local maximum. This variational problem is often simplified by assuming small deflection angles and approximating [41, 42, 101, 3].

The refractive index field can be related to the density of the fire reaction products, and through density can be related to temperature and pressure. The Gladstone-Dale equation codifies the simple empirical relationship:

$$(n-1) = k\rho, \tag{2.4}$$

where ρ is the density and k is a constant that varies across different species of gases. This equation can be shown to be mathematically equivalent to a model where molecules are packed tightly into some fraction of the volume, and the rest of the volume is empty [120]. The Lorenz-Lorentz equation,

$$\frac{n^2-1}{n^2+2} = k\rho, \tag{2.5}$$

is slightly more precise, and can be justified using arguments from electromagnetic theory. The refractive index is also dependent on wavelength and molecular composition.

2.1.3 Fire as Gas

Fire can also be examined from the perspective of gases, because the high temperatures involved guarantee that the reaction products will be generated in the gas phase. We can use insights from chemistry to relate the density of the gases involved in fire with the temperature T and pressure P. According to the ideal gas law,

$$Pm = \rho RT,$$

where m is the molar mass of the gas and R is the universal gas constant. The ideal gas law is a steady state equation, therefore applying it in dynamic settings such as fire is only an approximation. Moreover, while most gases are close to ideal at ambient temperatures, this formula is only approximate for fire because deviations are known to increase with temperature. From the Gladstone-Dale equation (2.4) we have that $(n-1) \propto T^{-1}$ at constant density and pressure.

As a gaseous medium, fire is also subject to the laws of fluid mechanics, however we postpone the discussion of fluid mechanics models until §3.1.3. While fluid mechanics models are at the heart of detailed simulation methods for fire, we are unaware of any work that exploits these dynamic models to obtain better measurements of the physical properties of fire.

2.2 Measurement Techniques

We describe several techniques, apart from regular photography, for directly recovering measurements from a fire. These basic measurements, including temperature and beam deflection, can be used to infer other quantities from the fire such as density. Note that we defer a discussion of tomography for combining multiple measurements into a single field until §4.1.4. The chemical composition of fire is usually determined by analyzing peaks across a broad infrared spectrum.

2.2.1 Temperature

The most straightforward methods for determining the temperature of a fire involve placing a probe directly in the flame. For temperatures below 300 °C, a regular thermometer is sufficient. Up to temperatures of 1500 °C, a more sophisticated version of a thermometer known as a thermocouple can be used. Thermocouples consist of a wire probe made

of two different metals. Because the metals expand at different rates with temperature, the voltage potential across the probe can be precisely calibrated to give temperature readings.

Two-wavelength pyrometry is an alternative method for estimating temperature by wholly optical means, and has the significant advantage of being non-intrusive. It involves recording the spectral radiance L for two different wavelengths λ_1 and λ_2. By manipulating Planck's equation (2.1) we obtain:

$$T = \frac{\frac{hc}{k_B}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)}{5\ln\left(\frac{\lambda_2}{\lambda_1}\right) + \ln\left(\frac{\xi_{\lambda_1}(T)}{\xi_{\lambda_2}(T)}\right) + \ln\left(\frac{L(T,\lambda_2)}{L(T,\lambda_1)}\right)},$$

where the emissivities for both readings are also required. However, the two-wavelength pyrometry method can only be applied inside the boundary of the visible flame, because blackbody theory fails when little or no light is emitted.

2.2.2 Deflection Measurements

The small amount of refraction present in flame can be exaggerated using a variety of special imaging techniques which modify brightness according to deflection. These methods have traditionally been used in a qualitative manner, and are particularly useful for visualizing non-luminous flames. Recent advances, however, have also allowed these deflection measurements to be made quantitatively.

One popular imaging technique for exaggerating deflection is *schlieren photography*, first developed in the early twentieth century to detect aberrations in lenses [36]. In fact, the word schlieren is German for streaks in an otherwise transparent region. Schlieren photography is contrast enhancing, and permits the visualization of fire without visible flame (see Figure 2.3). This is important for oil well fires, for example, which are only visible to the naked eye through the subtle shimmering of the background caused by refraction.

The schlieren apparatus, as shown in Figure 2.4, consists of a light source positioned behind the flame, a knife-edge filter (half of it opaque) between the flame and the camera, and several auxiliary lenses. The knife-edge filter deprives the image of light in downward deflection zones, removing specific harmonics from the diffraction pattern, but allowing unimpeded passage to the rest of the light. In this way, the image intensity is related to the deflection due to a field of varying refractive index and the focal length. The

Figure 2.3: Photograph taken using the schlieren method. Courtesy the University of Tasmania.

apparatus for schlieren photography is typically calibrated using a known light source and a prism of known refraction angle.

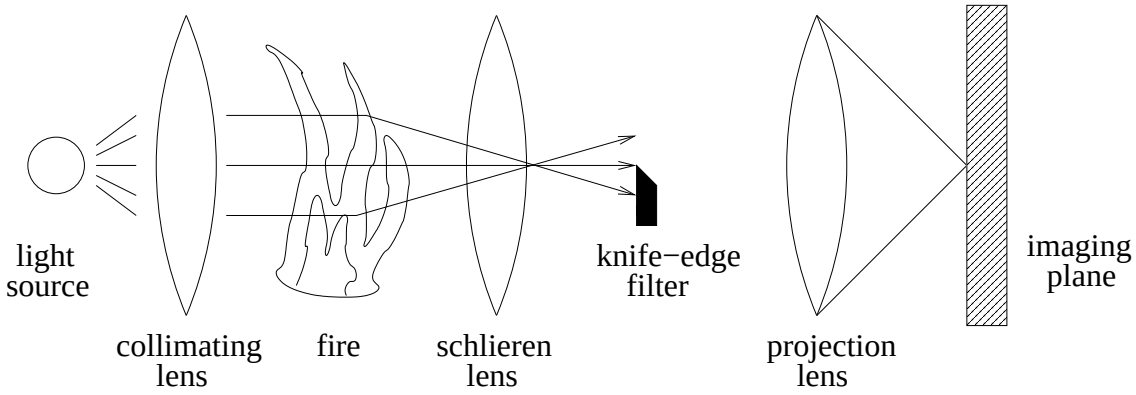


Figure 2.4: Apparatus for schlieren photography. Adapted from [120].

The newer rainbow schlieren method is very similar except that a computer-generated colour gradient filter is used instead of a knife-edge filter [3]. The coloured filter enables hue in the image to be precisely related to the amount of deflection.

Shadow photography is related, simpler technique for visualizing deflection. It also involves a light source, but uses no special filters for removing specific harmonics. The shadow method instead relies on shadows cast by the flame due to the deflection of

rays in regions of non-uniform density. The inhomogeneity of the flame again acts to redistribute the light, although in this case, the relative illumination change depends on the derivative of the amount of deflection. The image produced by the shadow method is harder to interpret, especially for turbulent flame, since only the darkened zones provide a representation of the disturbance geometry. Brightened zones cannot be related so easily to deflection, and the shadow method has not been used quantitatively for this reason.

Diffraction methods can also be used to obtain deflection measurements by examining the fringes of a diffraction pattern projected through a fine grating. More specialized methods rely on detecting fluorescence and resonance scattering at non-visible wavelengths [120].

2.3 Image Formation Models of Fire

We present several different image formation models of fire in order of increasing complexity, and describe the relationships among them. For ease of presentation, we ignore the contribution of the background intensity. In practice, this contribution is easy to incorporate as an additional intensity term attenuated by the absorbance of the fire and possibly warped by refraction. Moreover, in many experimental situations, the background can be controlled so that its intensity is negligible.

2.3.1 Soot Volume Model

Most liquids and solids burn with luminous diffusion flames (see §2.4). The characteristic yellow luminosity is a net result of emission from minute carbonaceous particles formed in the flame, mainly on the fuel side of the reaction [36]. These particles typically measure 10–100 nm in diameter, and approximate tiny blackbodies with continuous emission spectra inside the flame. While the particles are often consumed by oxidative burning, they may react further or cool to yield smoke.

The theory explaining the soot volume model predicts that for particle diameters less than the radiation wavelength, the emissivity should be proportional to the density of the soot [36]. For visible light, whose wavelength is in the range of 400–700 nm, the particle diameters are small enough to satisfy this requirement.

The soot volume model is analogous to regular spectroscopy, but with the density of

the medium ρ controlling the emission as well as the absorption. By applying Kirchoff's Law, equating emissivity and absorptivity, we are in essence interpreting the emission from a flame as the degree to which the same soot density would have absorbed light. This insight allows us to make direct use of the Beer-Lambert Law from Equation (2.2) giving the emissivity ξ_λ as:

$$\xi_\lambda = 1 - \exp\left(-\sigma_t \int_0^L \rho(s)\,\mathrm{d}s\right), \tag{2.6}$$

where σ_t is the extinction cross-section and L is the path length as in §2.1.1. While the soot volume model is formulated in terms of emissivities, the emitted radiance is dependent on both emissivity and temperature. Thus, temperature variations within the fire must somehow be taken into account if the soot density is to be accurately reconstructed. The simplest approach is to bypass the problem entirely by assuming that the temperature of the fire is constant.

2.3.2 Radiative Transfer Model with Simplifications

There are more comprehensive theories of light transport in participating media as well, including the well-known radiative transfer model [37, 96]. In its general form, the radiative transfer model explains emission, absorption, and (single and multiple) scattering. For a clear introduction to the subject building from basic ideas, refer to the tutorial paper by Max presented to the computer graphics community [70].

The transparency over a given ray is defined in the radiative transfer model as a familiar line integral over the density field $\rho(s)$:

$$\tau(x) = \exp\left(-\sigma_t \int_0^x \rho(s)\,\mathrm{d}s\right). \tag{2.7}$$

The ray is parameterized so that x represents the distance from the eye into the medium, and the extinction cross-section σ_t is a constant relating the density and the transparency.

In the radiative transfer model, the source function $J(s)$ represents the total effect of emission and inscatter, less any absorption and outscatter. The source function is often decomposed as $J(s) = \Omega S\{I(s)\} + (1-\Omega)Q(s)$, where the albedo $0 \leq \Omega \leq 1$ represents the fraction of scattered versus absorbed light. The scattering events are given by some functional form $S\{I(s)\}$, where the first scatter is usually described in terms of a phase function giving the redistribution of light over a sphere. Subsequent scatters are frequently approximated as a diffusion process for computational efficiency

[107, 57]. The emission $Q(s)$ is a combination of inscatter and self-emission. Putting everything together, the intensity along a ray $I(x)$ predicted by the radiative transfer model can be written as:

$$I(x) = \int_0^x \tau(s)\sigma_t\rho(s)J(s)\,\mathrm{d}s, \tag{2.8}$$

where the ray is again parameterized by x and σ_t is the same extinction cross-section used to define the transparency τ. To simplify the notation, we have dropped the dependence on wavelength.

In nearly all image formation models of fire, the albedo is assumed to be zero, meaning that scattering is assumed to be negligible [101, 107, 31, 124, 3]. This approximation, also known as total emission, performs quite well for (smokeless) fire because self-emission from glowing soot particles dominates the observed intensity. Moreover, the little scattering actually present in fire can often be accounted for as an additional contribution to the self-emission term. While total emission is a less appropriate assumption for smoky fires, the phenomenon of smoke is typically modelled separately in both the computer graphics and the combustion literature [107, 8]. Under the total emission assumption, we have that $J(s) = Q(s)$, where the emission $Q(s)$ consists solely of self-emission.

In general, self-emission can vary over the fire independent of the density. Some simulation models of fire calculate self-emission $Q(s)$ according to temperature $T(s)$ using Planck's equation for ideal blackbodies (2.1), converting the full spectral distribution into RGB values for display [107, 82]. More sophisticated simulation models account for the varying spectra of different chemical species present in the reaction products [84, 95].

Further simplification occurs if self-emission (per unit mass) is assumed to be constant $Q(s) = Q_0$. Under this assumption, both absorbance and emission are functions of a single density field, as in the soot volume model. With the self-emission proportional to density, the radiative transfer equation (2.8) simplifies greatly to give:

$$I(x) = (1 - \tau(x))\,Q_0. \tag{2.9}$$

One caveat is that the effect of temperature is not properly considered in this simplified model. Any density field reconstructed using this model will represent the combined effect of mass density and temperature, rather than the true density in isolation. Note that the form of this model is identical to the soot volume model given in Equation (2.6) except for the constant factor of Q_0 at the end.

2.3.3 More Complex Models

A more complex image formation model of fire was presented by Schwarz for the purposes of reconstruction [101]. His model involves the combination of three scalar fields: emission coefficient E, extinction coefficient μ, and refractive index n. Apart from modelling refraction effects, his model is identical to the zero-albedo radiative transfer model in Equation (2.8) in which self-emission and the absorbance are decoupled. In relation to our previous definitions, we have that $E(s) = Q(s)\rho(s)$ and $\mu(s) = \sigma_t\rho(s)$.

Schwarz created an sophisticated apparatus with a laser light source and schlieren filters. By combining various images with and without the fire, the laser, and the filter, he was able to isolate and reconstruct the different scalar fields. For instance, the extinction field can be reconstructed from the combination of three images:

$$\ln\left(\frac{I_{laser}}{I_{laser,flame} - I_{flame}}\right) = \int_0^L \mu(s)\,\mathrm{d}s.$$

Once the extinction field is known, the emission field can then be reconstructed using a formula analogous to Equation (2.8):

$$I_{flame} = \int_0^L E(s)\exp\left(-\int_s^L \mu(t)\,\mathrm{d}t\right)\,\mathrm{d}s,$$

Reconstructing the refractive index field is more challenging. For luminous flames, this involves combining six images and solving a more difficult tomography problem that no longer consists of simple straight-line integrals. Schwarz reconstructed the refractive index field for both axis-symmetric and non-luminous flames.

2.4 Qualitative Fire Dynamics

We discuss several important aspects of fire dynamics in a qualitative manner. For a more detailed treatment, see [120, 36]. The factors controlling flame motion and propagation are numerous. They include the buoyancy of hot air, gravity, pressure gradients, and the burning of the fuel.

Flames can be separated into two broad categories: premixed flames and diffusion flames. Premixed flames are generated when a mixture of fuel and oxygen is ignited. These flames burn rapidly and the final burnt state is completely determined by the composition of the fuel. Diffusion flames, on the other hand, are characterized by an initial separation of reactants. The relatively slow mixing of the reactants acts to control

the rate of combustion, and the final burnt state depends on the nature of the particular mixing process.

The mixing process for diffusion flames can be homogeneous, as with the outer cone of flame from a gas burner with constant flow. The mixing process can also be heterogeneous if mixing is determined by heat and mass transfer. This is the situation for the combustion of a liquid fuel droplet or burning along a wick.

We can also characterize the motion of flames as turbulent or steady. Turbulent flames fluctuate randomly (on a local scale) in velocity and pressure, producing characteristic swirling eddies (see §3.1). When turbulence is dominated by buoyancy effects, it is often referred to as flickering [36]. Steady flames, as the name suggests, are extremely controlled. One class of steady flames is laminar flames, in which gases flow mainly in parallel layers with different velocities.

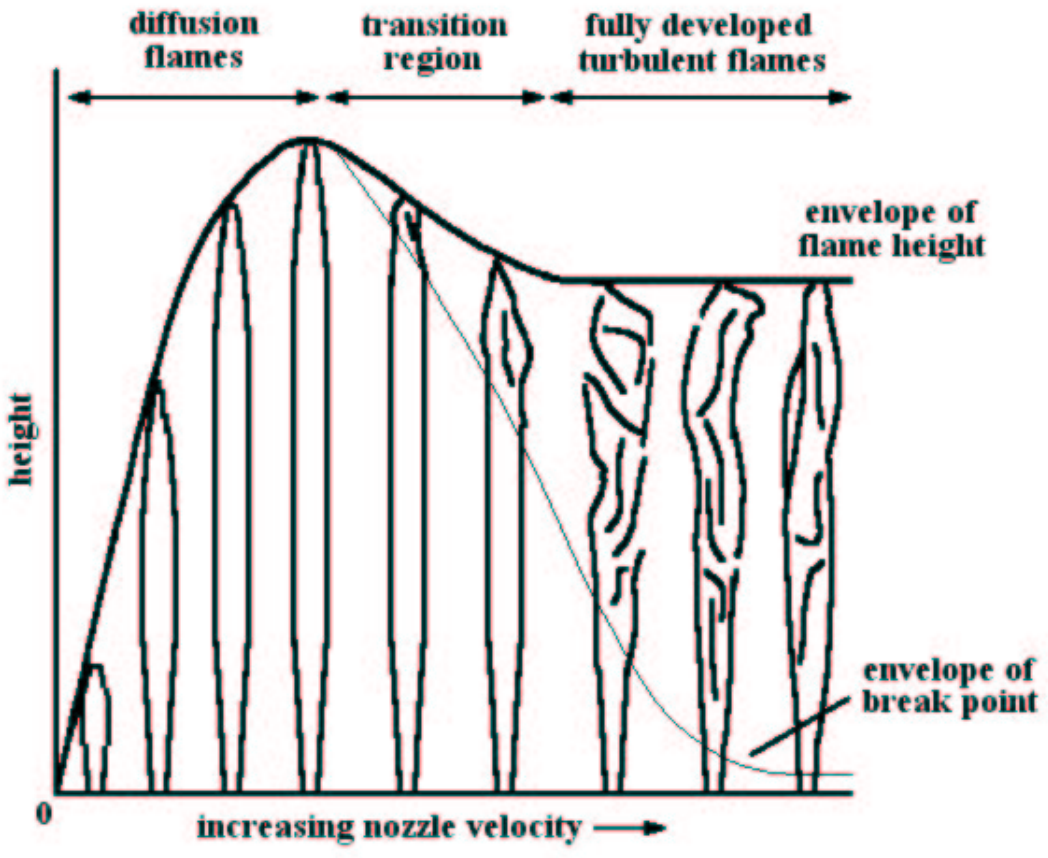


Figure 2.5: Variation of flame shape with nozzle velocity. Adapted from [120].

The relationship between turbulence and nozzle velocity for a fuel jet flame is illustrated in Figure 2.5. Flame height is shown to increase without significant turbulence up to some maximum velocity. Then, with increasing nozzle velocity, the upper portion of the flame becomes increasingly turbulent and the flame height levels off until turbulence completely dominates.

Another important combustion process is the quenching of flames. Quenching can occur by the proximity of relatively cool barriers, diffusion, or suspended particles (as in the halide fire extinguishers that protect expensive computer equipment). Because quenching removes combustible reactants from the system, it is considered to be a non-steady state process.

To give a concrete example bridging several of these concepts, consider a typical laboratory burner that burns fuel-rich mixtures in air. These flames are steady, consisting of both a premixed inner cone and diffusion flame outer cone. The homogeneous mixing rate in the outer cone is determined by the constant gas flow and the diffusion of the hot reaction products.

2.5 Summary

Fire can be viewed from a variety of different perspectives, and many of its physical properties are interrelated. Fire is at once a source of radiation, a gas obeying the laws of fluid mechanics, and a transparent medium modulating light through a varying refractive index field. We have concentrated on the image formation model of fire because our main interest is the visual reconstruction of fire and those physical properties that can be measured optically.

Most image formation models characterize fire as an emissive but highly transparent medium and assume that fire is non-scattering. In the absence of significant smoke or saturated images, these assumptions are reasonable. The refractive properties of fire are less relevant for regular image formation, however an appropriate apparatus can be used to exaggerate beam deflection dramatically and help visualize the contours of othewise invisible flames.

Chapter 3

Fire in Computer Graphics

> Persons grouped around a fire or candle for warmth or light are less able to pursue independent thoughts, or even tasks, than people supplied with electric light.
>
> *Understanding Media* (1964)
> Marshall McLuhan

From the perspective of computer graphics, the main concerns in modelling physical phenomena such as fire are visual realism and efficiency. This stands in sharp contrast to engineering and the physical sciences, where the goal of simulation is to accurately compute various physical quantities to specified tolerances. Of course, visual realism is enhanced by physically accurate modelling, but sometimes we can find good compromises where visually negligible approximations can be traded for significant gains in efficiency.

This chapter reviews a number of methods for modelling and rendering fire, adapted for the purposes of computer graphics. The bottleneck in this process is modelling the fire with sufficient accuracy and resolution. We will see later that graphics hardware already allows good quality rendering at interactive rates.

In fact, the scope of this chapter is somewhat broader than the modelling and rendering of fire. We consider all manner of gaseous phenomena, since the techniques developed for general turbulent gases can often be used directly for the purposes of fire. Until recently, fire had not been afforded its own separate treatment in the computer graphics literature. Instead fire was regarded as a special case of more general gaseous phenomena [98, 107, 45] or an interesting application of various techniques like particle systems [93]. Any rendering techniques not directly applicable to fire, for example multiple scattering in rendering smoke particles, will be highlighted.

3.1 Modelling Fire for Computer Graphics

Visually realistic models of fire need to capture the chaotic motion of turbulent gases. So far, the major modelling approaches for turbulence have been combining various procedural elements, characterizing the global statistics of turbulence, and solving simplified versions of the fluid mechanics equations. Though some of the techniques described in this section were originally developed for broader classes of gaseous phenomena (including fog, smoke, and steam), their contributions are important for modelling fire as well.

Another aspect of fire that must be modelled to achieve realism is the combustion of fuel. To this point, combustion models have been of secondary importance for computer graphics, perhaps because most applications do not involve observing fire for long enough for a casual viewer to detect problems with the combustion model.

3.1.1 Procedural Models of Fire

Procedural models for the simulation of fire are based on heuristics describing the rules that govern the dynamics of simple primitives. These methods are not strictly based on physical principles, though tuned appropriately they can nonetheless generate interesting visual effects. Furthermore, procedural models often have the advantage of being computationally efficient. The main drawback of these methods is that controlling the phenomena can be difficult, with many brittle parameters to tweak in order to achieve reasonable results. Moreover, extending the models to obtain new effects or to interact with other elements in the scene is often impossible.

Chiba *et al.* developed a 2D model of fire based on a procedurally defined turbulence field [25]. They represented turbulence using vortices as primitive elements, and visualized the motion of the flame using motion-blurred particle tracers. The vortices are described by a simple empirical model (Rankine-type vortices), in which an inner core behaves like a rigid rotated body, and the velocity outside the core decays with the inverse distance from the centre. Heuristic rules were also defined for the random appearance, growth, and diffusion of the vortices.

This method was later extended to 3D, with vortices modelled using a rotating inner core in the shape of a double cone [114]. The results of both methods are visually acceptable, and a range of different types of fire can be simulated by manipulating different variables. These variables include the number of vortices, the upward force for mimicking buoyancy, and the tendency for the vortices to attract. However, overall control is still

rather limited and requires many parameters to be adjusted.

Another procedural approach for fire simulation, based on cellular automata, was described by Takai, Ecchu, and Takai [115]. Cells in a 3D grid that do not contain obstacles can be occupied by a single fire particle, labelled with a mass and a discretized velocity. Randomized transition rules control the movement of the particles, localized to small neighbourhoods. Then, each time step involves solving a collection of small optimization problems, handling the collisions between particles and obstacles as well as possible, while satisfying constraints such as the requirement that no more than one particle occupy a cell. The approach is fairly efficient, and the fact that transition rules are local allows good parallelism. While cellular automata of this kind are usually fast, their behaviour is especially sensitive to the specific transition rules; changing the rules even slightly can break the entire system. However, impressive results recently been demonstrated using an extension of these techniques for simulating 3D clouds [33].

Lamorlette and Foster define a procedural model of fire based on curve structures that also incorporates a number physically realistic elements [65]. Their goal is to give the animator a high degree of control while still producing visually plausible results. The overall geometric structure of their fire model is determined by radial basis functions over interpolating B-splines or by hand-drawn flame profiles swept out rotationally. These curves are warped by turbulent wind fields (§3.1.2) and interact with solid obstacles using simplified fluid mechanics (§3.1.3). The flames are further divided into different procedural zones. Above a persistent base of flame, a zone of intermittent flickering periodically releases short-lived buoyant flares. This model produces realistic visual results for certain types of fire and was successfully integrated into a production film environment.

3.1.2 Spectral Synthesis of Turbulence

One established approach for characterizing a turbulent field is through a global, statistical description. This approach allows plausible turbulence to be generated efficiently, while avoiding the intensive computation associated with a full simulation of fluid mechanics. Turbulence in its steady-state is known to be well-characterized by the Kolgomorov energy spectrum [98, 110]:

$$E(k) = \begin{cases} 0 & \text{if } k < k_{\text{inertial}} \\ 1.5\,\varepsilon^{2/3} k^{-5/3} & \text{otherwise} \end{cases}$$

where the largest scale eddies below frequency k_{inertial} are assumed to play no role in the turbulent motion. The hypothesis here is that energy is only lost to friction by eddies at the smallest scales. At intermediate scales the energy is redistributed at a constant rate ε in a cascade to the higher frequencies. Note that models of this form have also been derived from experimental data [86].

The precursor to the spectral synthesis of turbulence in computer graphics was Perlin's heuristic method for generating turbulence [87]. Perlin found that turbulent-like fields could be created by summing his eponymous noise function with a series of scaled-down copies of it. This "fractal" sum is similar in spirit to the continuous $k^{-5/3}$ falloff found in the Kolgomorov energy spectrum.

Sakas used the spectral synthesis approach for simulating clouds by directly synthesizing turbulent density fields in the spatial domain [98]. Taking the inverse Fourier transformation of the Kolgomorov energy spectrum, he obtained plausible turbulence in 4D. The last dimension of this space was treated as time, in order to produce smooth animations of 3D turbulence. Sakas also derived a technique that allowed him to apply simple translations in the spatial domain (a global wind vector) to the turbulence.

The spectral synthesis method proposed by Stam and Fiume overcomes a number of limitations in Sakas's work [109]. By generating a turbulent velocity field instead of generating the density field directly, their method eliminates periodicity artefacts in the density field, allows the animator to incorporate arbitrary large-scale wind fields, and enables particles convected by the velocity field to interact with obstacles in a realistic way. Stam and Fiume also suggest modelling the dependency on temporal frequency separately. With this modification, smaller eddies can be made to be less correlated in time, thus spinning and flowing faster.

3.1.3 Fluid Mechanics Models

The most physically realistic models of fire (and other gases) for computer graphics are those based on principles from fluid mechanics. This approach eases the burden of the animator considerably, ensuring visually plausible results, but restricts control over the fire. Some effort can be focused on setting boundary conditions and imposing other constraints, however achieving direct control for dramatic purposes is very difficult using purely physically based models. Fire as a background element is readily modelled using these techniques, but precise control demands a hybrid method combining physical

realism with user-specified procedural elements [65].

The general approach of fluid mechanics models of fire is to discretize the domain and use finite element methods to solve a simplified form of the Navier-Stokes equations. However, this method is computationally intensive, scaling by at least $O(n^3)$ with the resolution of the simulation. Furthermore, the solver must be implemented in a relatively sophisticated manner to avoid common problems with instability.

The numerical methods become much simpler with the assumptions that the fluids are incompressible and of constant density [45, 108, 6, 43]. These assumptions are justified for many types of gases with relatively low velocity and pressure, although extreme effects like shock waves from powerful jets are not modelled. For fire, the assumption of incompressibility is somewhat more problematic because the expansion of the fuel as it reacts to form hot products is visually important. For reasons of efficiency, models of fire typically retain the incompressibility assumption but account for the expansion of hot gases as the result of external forces [82, 45].

Another simplifying assumption is that the fluids are subject to negligible viscous dissipation [6, 43]. This is a good assumption for gases, rather than liquids, especially where the damping effects of the numerical methods are large enough to mask the effects of viscosity.

By assuming no viscosity and arbitrarily setting the density to one, the Navier-Stokes equations for incompressible fluid flow simplify to:

$$\nabla \cdot \mathbf{u} = 0 \tag{3.1}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -(\mathbf{u} \cdot \nabla)\mathbf{u} - \nabla p + \mathbf{f}, \tag{3.2}$$

where $\mathbf{u}$ is the velocity, p is the pressure, and $\mathbf{f}$ accounts for external forces. Equations (3.1) and (3.2) ensure that mass and momentum are conserved respectively. As usual, boundary conditions can be specified either directly or in terms of constraints on the gradients. Foster and Metaxas [45] enumerate a nice collection of boundary conditions for achieving various special effects. These include solid obstacles with smooth and rough surfaces, steaming hot soup, and open windows.

Two additional scalar fields, temperature and local density, are also modelled. These quantities are advected at every time step by the velocity, and for more accuracy the diffusion of heat to adjacent regions can be modelled as well [45]. For even finer control, Nguyen, Fedkiw, and Jensen suggest allowing the animator to create an explicit profile representing the progression of fuel temperature over time [82]. This profile can then

be mapped to a fluid element according to how much time has elapsed since it was first generated in the fuel zone. One physical phenomenon dependent on both temperature and density is thermal buoyancy, the tendency of hot gases to rise and heavy smoke to fall. Thermal buoyancy is modelled as an external force directly proportional to density and temperature.

Foster and Metaxas were among the first to successfully apply fluid mechanics to 3D computer graphics [45]. Their model captured the rolling, billowing motion of gases as well as interesting mixing effects between gases of different temperatures. However, because their method was computationally intensive and suffered from problems of instability, the results were necessarily low-resolution.

Stam ignored the effect of temperature and proposed an efficient solution method for incompressible fluids with guaranteed stability properties [108]. In particular, by assuming (non-physical) wrap-around boundaries, the system can be solved in a particularly elegant fashion, using only a particle tracer and the Fast Fourier Transform (FFT). While Stam's stabilized method has the advantage of allowing arbitrarily large time steps, on relatively coarse grids the method is noticeably overstabilized; the smaller vortices become damped out by a phenomenon known as numerical dissipation. Arquès *et al.* pursued a very similar avenue of research [6]. Their group proposed an implicit integration scheme for stabilizing an incompressible fluid solver based on geometric insights.

In their recent work on the simulation of smoke, Fedkiw, Stam, and Jensen used the method of vorticity confinement to address the numerical dissipation problems found in fluid solvers incorporating stabilization [43]. The idea of vorticity confinement, borrowed from the engineering community, is to add back small-scale turbulent detail through a forcing term, but to do so only in the physically correct locations. Another innovation of theirs for reducing numerical dissipation is a new higher-order interpolation scheme free from overshooting.

A new fluid mechanics model specialized to fire was suggested Nguyen, Fedkiw, and Jensen [82]. Their model involves two distinct zones, a premixed inner core and outer diffusion flames (see §2.4), separated by an implicit surface. This surface, defined using level sets, serves as a reaction front between the premixed fuel and the reaction products. Each side of the reaction front is assigned a different constant density so that hot gaseous products will expand in keeping with conservation of mass and momentum. This approximation is fairly accurate for premixed gaseous fuel jets, but with the proper choice of parameters, larger diffusion flames can be reproduced acceptably as well. For solid fuels,

the reaction front can be chosen as the surface of the fuel, but the results are far less convincing.

3.1.4 Models of Combustion

Realistic models of fire should also be able to interact plausibly with the environment through the process of combustion. Accurately modelling combustion involves the consumption of fuel, the radiation of heat, and the ignition, spreading, and extinguishing of flames. Many of the computer graphics models we have considered so far encompass gaseous phenomena in general, but even those specialized to fire sometimes ignore this aspect of the modelling altogether [62] or approach the modelling of combustion in a completely *ad hoc* manner [25, 115].

One heuristic approach to combustion modelling involves decomposing the burnable objects into voxels and labelling each voxel with variables indicating the presence or absence of burning, the temperature, and the quantity of fuel remaining. Then, hard temperature thresholds can be set for ignition and extinguishing, and heat transfer can be modelled as diffusion to adjacent cells [25, 115]. Heuristic models of combustion have also been defined over polygonal meshes, where the evolution of the flame front is given in procedural terms [9, 65]

A somewhat more sophisticated and physically based approach was suggested by Stam and Fiume [110]. They invoked the Arrhenius formula for flame production, which is based on an activation temperature and the assumption that the oxidants are maintained at a constant concentration. Heat transfer was then realized by calculating thermal radiation using the blackbody approximation, and using radiosity-style shooting operations to send heat to nearby patches.

3.2 Rendering Fire

Fire is a quintessentially volumetric phenomenon, visually characterized by its emissivity and transparency. Rendering fire demands some form of volume rendering, with special attention paid to the image formation model that is being assumed (see §2.3). We examine general volume rendering techniques, the use of graphics hardware for rendering fire at interactive rates, and more complicated models that treat the fire as a source of illumination and capture mirage-like refraction effects. These rendering techniques

are described in the context of more general techniques for rendering smoke and other gases, even though certain effects like scattering are ignored by nearly all image formation models of fire.

3.2.1 Scalar Fields Describing Fire

All of the rendering techniques described in this section operate on the same kind of input: one or more 3D scalar fields describing various salient characteristics of the fire. It is the role of the modelling process is to produce 3D distributions for such variables as density, emissivity, optical thickness, refractive index, temperature, and chemical composition. Any image formation model of fire makes assumptions about which of these variables are important and the relationships between them. For example, under the soot volume model of fire, the emissivity and temperature fields are needed for rendering.

These scalar fields are calculated differently depending on how the fire is modelled. For some of the fluid mechanics models based on finite-element methods, the density and temperature fields are calculated directly and may be used without further manipulation [108, 6, 43]. The density field may also be represented implicitly in terms of basis functions. Stam and Fiume use Gaussian blobs, warped by advection and diffusion, as a primitive for the fire density field [110].

Other approaches using fluid mechanics do not model the density field explicitly. Instead, a number of massless particles can be introduced into the system to be advected by the turbulent velocity field. The density of the fire can then be approximated by the concentration of particles in each voxel [45]. A related approach involving particle tracers is presented in the work on procedurally defined vortex fields [25, 114]. There, many fewer particles are needed because the particles are blurred over both space (as the centres of radial basis functions) and time.

3.2.2 Scientific Visualization

The purpose of scientific visualization applications is to convey information effectively, rather than to depict the phenomena being modelled in a realistic way. Rendering fire for scientific visualization involves determining what information is most important for the domain, abstracting this information, and employing an eclectic range of rendering techniques.

One example of such an application is the system presented by McCormick for the visualization of wildfires [71]. He used a number of different strategies to render the fire data, including pseudocolour, distinctive textures for labelling categories of burning land, and transparent polygonal isocontours. The dynamics of the wildfire were visualized using a video recorder style interface to control the animation and by overlaying 3D arrows to depict the velocity field. Similar techniques were used in a system developed for the simulation of fire in building environments [22].

3.2.3 Direct Volume Rendering

The main approach for rendering fire is direct volume rendering. In other words, we consider the volume directly, without introducing intermediary representations like polygonal approximations of isocontours.

Recall that the general radiative transfer model of fire in Equation (2.8) involves three different terms: the density field ρ, the extinction cross-section σ_t (a constant relating density and transparency), and a source function J (accounting for emission and other light sources). Note that scattering is always assumed to be negligible in fire. We restate this image formation model:

$$I(x) = \int_0^x \tau(s)\sigma_t\rho(s)J(s)\,\mathrm{d}s, \tag{3.3}$$

where I is the intensity computed along a ray, τ is the transparency from Equation (2.7), and the contribution of any background intensity is ignored for now. Because the source function is decoupled from the density field, it can be made to depend on other variables such as temperature.

Since the integral in Equation (3.3) is difficult to evaluate analytically, the usual approximation is to break it into n intervals and assume that the source function is piecewise constant over these intervals. For a voxelized representation, a natural choice for the intervals are the voxels themselves, but these intervals may be determined otherwise for different primitives like warped blobs. This approximation gives the usual compositing equation [110]:

$$I = \sum_{i=1}^{n} \left(J_i \left(1 - \tau_i\right) \prod_{j=1}^{i-1} \tau_j \right), \tag{3.4}$$

where J_i is the source function constant over the interval i.

A somewhat easier case is the simplified radiative transfer model of fire, in which emission is assumed to be proportional to density. Under this assumption, the source

function is constant $J = Q_0$, and Equation (3.3) simplifies to give $I = (1 - \tau)\, Q_0$. Thus, image intensity is explained in terms of line integrals over the density field.

The most straightforward method of volume rendering is a modified form ray tracing for evaluating these line integrals. Digital differential analyzer (DDA) methods, analogous to Bresenham line drawing in 2D, are used to step through the 3D voxel grid in an efficient manner, accumulating intensity along the way. This intensity integral for the fire can be blended appropriately with the background intensity to give the effect of transparency. To be more precise, the integration should take place over the point spread function centered on a given image pixel, but for relatively long focal lengths the added accuracy is almost never worth the computational effort.

For special cases where the volume is represented as cylindrical rings [95] or spherical shells [83] the integration over rays can be accelerated. When the density field is defined implicitly in terms of primitives such as Gaussian blobs, constructing lookup tables can help deliver additional speed [110].

Faster and more sophisticated methods of volume rendering have also been developed. One simple optimization to accelerate ray tracing is terminating early if the transparency ever becomes zero. Another volume rendering technique, known as splatting [121, 67], avoids the expensive 3D interpolation required by ray tracing. In splatting, radially symmetric kernels are placed at the centre of each voxel. These kernels are pre-integrated, with their different footprints for the resulting 2D image stored in a table. The contributions of the voxels are then accumulated directly in the image, with all calculations are performed in 2D. Newer splatting methods have also been proposed to address common artefacts such as blurriness [76].

Stam suggested another interesting approach to rendering turbulent volume fields [106]. Under his stochastic rendering framework, the scene to be rendered is represented as a random field with known statistics. These statistics are then shifted in a systematic fashion from the modelling phase to the rendering phase. This allows relatively simple primitives to be used in modelling, but more complex images to be obtained from the rendering. Stam derived analytic expressions for the random statistics of the transparency function and developed good approximations for describing the source function, which cannot be solved analytically in this way.

3.2.4 Hardware Acceleration Techniques

Certain techniques that exploit fast graphics hardware enable the rendering of fire at interactive rates. Perhaps the simplest technique for rendering fire quickly involves applying movie loops as textures to billboards [73]. The set of textures can be generated using spectral turbulence methods, or by extracting pieces from real images of fire. Unfortunately, the results obtained using this simple method are unconvincing.

King, Crawfis, and Reid combined this texture-based approach with traditional volume rendering using splatting [62]. They used a coarse voxel grid, where each voxel was assigned an index into the set of textures. By manipulating the texture indices as if they were particles, they gave the impression of smooth motion. When performing splatting, the kernels were modulated with the appropriate texture corresponding to that voxel. In contrast to the billboard method, their method is effective even when the viewpoint is inside or very close to the fire.

Hardware methods have long been used for accelerating standard volume rendering [59, 34, 68]. These approaches require at least one additional frame buffer (alpha buffer) to efficiently accumulate the intensity using the compositing Equation (3.4). In one such method [43], volume slices in the direction most closely oriented to the camera are rendered from front-to-back as semi-transparent quadrilaterals, appropriately textured by the densities and radiances of the voxels in that slice.

More sophisticated effects are also amenable to hardware acceleration. Self-shadowing effects, for example, can be obtained by adding an additional pass for each light source [108, 43]. Ray tracing, again made efficient using a DDA method, can be used to step through the voxel grid and update the radiances appropriately. The transparency of the ray is made to diminish as it traverses the density of the volume, creating the shadowing effect. While self-shadowing is important for rendering clouds and smoke, it is not as significant for rendering fire because fire is usually far brighter than anything else in the scene.

This multi-pass approach using graphics hardware was also main basic idea used by Dobashi *et al.* for rendering clouds [33]. He extended the method to render shafts of light by interleaving the rendering of volume layers with spherical shells created from transparent, texture-mapped polygons.

3.2.5 Illumination from Fire

For added realism, we would like the fires we are modelling to act as light sources and cast illumination into the scene. Takahashi, Takahashi, and Chiba proposed a method for doing this similar to volume rendering using splatting [114]. By assuming that the viewpoint and the objects in the scene (other than the fire) are static, they constructed an efficient rendering method using footprints precomputed for each voxel over a coarsened version of the voxel grid.

Note that slower, high quality methods are also available for rendering scenes containing fire. For example, radiosity methods [110] and photon mapping [43] have been used extensively for rendering gaseous phenomena. Since both methods model the global propagation of light in a scene, they capture more subtle effects such as multiple scattering (which is less relevant for smokeless fire) and illumination cast from flames.

3.2.6 Rendering Refractive Properties of Fire

In the context of computer graphics, refraction has long been incorporated into ray tracers, but only abrupt changes in refractive index between different media are typically modelled. The first graphics techniques to model continuously varying refractive index fields were those addressing the rendering of mirages. Berger, Trout, and Levit describe a model of mirages where the temperature above the ground drops off rapidly, leading to parabolic rays [11]. To approximate this effect, they build a transparent mirage object using a stack of layers with constant refractive indices, then ray trace as usual. Stam modelled the refractive index field as a superposition of blobs in a turbulent air field and calculated the perturbations due to these blobs [111]. He then used his stochastic ray tracing method to shift the random turbulence from the refractive index field to the rendering step. The total effect is realistic deformations and shimmering.

Oddly enough, the techniques developed for rendering mirages and heat waves seem to have been largely ignored in rendering fire for computer graphics. Arquès *et al.* presented the only fire known rendering approach so far to incorporate varying refractive index [6]. They use the Lorenz-Lorentz equation (2.5) to establish a connection between simulated temperature and refractive index, noting that density varies inversely with temperature. When they ray trace, rays are deflected accordingly at the borders of each voxel.

3.2.7 Colour of Fire

In rendering fire, the flames are often coloured according to some hand-coded heuristic function. For flames generated using a two-dimensional profile, one suggestion is to map colours from a reference photograph onto that shape [65]. More commonly, plausible looking colour maps can be devised corresponding to a range of temperatures or densities [115, 114, 6, 71]. For a typical campfire, one might devise such a colour map by smoothly blending across the sequence grey-red-orange-yellow-white, as the flame goes from cold to hot. The grey at the beginning of this sequence is a trick to obtain the appearance smokiness in the coolest regions. For different types of fire (blue natural gas flame) or other phenomena rendered in this way (clouds, smoke) different colour maps may apply.

While the results obtained using a carefully chosen colour map can be visually convincing, a more rigorous approach is also possible. The Planck equation for blackbodies (2.1), sometimes incorporating a single empirical scaling factor, has also be used to obtain colour spectra as a function of temperature [107, 82]. Rushmeier and Hamins used a software tool called RadCal, developed by the combustion community, to estimate the emitted spectrum based on the temperature and the estimated concentrations of different chemical species (water vapour, carbon monoxide) [95]. For maximum accuracy, the full spectra resulting from these calculations was convolved with the CIE basis functions and then converted to RGB.

In reproducing the colour of fire, another important consideration is the adaptation of our eyes to the fire spectrum. Fire is typically the brightest element in a scene, and because of the limited dynamic range, only very bright background objects can be seen through the fire. One suggestion is to map the spectrum for the maximum temperature present in the fire to the white point of the cone responsivities of the eye [82]. While our eyes adapt naturally to real fire over time, for computer graphics applications this effect must be simulated because typical display devices have a very limited maximum brightness.

3.3 Summary

Rendering convincing fire for computer graphics requires simulating the dynamics of flame with visual (if not physical) accuracy. While current models of combustion are still primitive, jets of flame have been simulated convincingly using simplified models of fluid

mechanics as well as more sophisticated procedural models. These models capture the essential dynamic behaviour of fire, including such features as diffusion, advection, buoyancy, and turbulence. For the purposes of animation, visual realism must be balanced carefully with precise control on the part of the animator.

The simulation of fire is the bottleneck in the fire rendering process. Given a three-dimensional representation of fire, images can be produced quickly using standard volume rendering techniques exploiting hardware acceleration. For more sophisticated results, techniques like photon mapping, which treat the fire as a true volumetric light source, can be used. Psychophysical effects like the adaptation of our eyes to bright flame can also enhance the realism of fire rendering.

Chapter 4

Vision and Fire

> The flame is not as bright to itself as it is to those it illuminates: so too the sage.
>
> *Human, All-Too-Human* (1878)
> Friedrich Nietzsche

As described in §2.3, a central concept is that fire can be modelled as an emissive but transparent medium. This makes handling transparency crucial for any attempt to reconstruct images of fire. While there is no general theory of handling transparent scenes in the vision literature, there have been attempts to account for object transparency in a variety of computer vision settings, from medical imaging and motion estimation to 3D reconstruction.

Transparency is easy to define in intuitive terms, though its precise definition is more slippery. What worsens things is that the the term transparency is frequently used to represent multiple effects in combination. In computer vision, transparency has been variously used to represent partial coverage of a pixel by a foreground object, the degree to which light from the background can pass through an object, uncertainty about the contents of a pixel, and different combinations of these [105, 128, 15].

Because the imaging model of fire involves transparency, the reconstruction of fire from synchronized images is closely related to tomographic reconstruction methods. Notably, tomographic reconstruction methods have been recently proposed in both the combustion literature [31, 8, 124, 41, 3, 125] and the atmospheric science literature [35, 90, 44]. While these tomographic reconstruction techniques constitute an advance over older measurement methods, tbey are still insufficient for our purposes. The main problem is that these reconstructions are too limited in both temporal and spatial resolution to be of interest for computer graphics or to validate detailed simulation models of fire. Further-

more, the techniques that yield the best results tend to require complicated apparatuses with dozens of synchronized cameras [101, 8]. This glut of hardware also runs contrary to our goals. Because our reconstruction methods (see Chapter 5) are designed to capture fine geometric detail with relatively few cameras, they should be valuable to the combustion community as well.

4.1 Computerized Tomography

By far the largest field of research that involves the imaging of transparent objects is computerized tomography (CT). While the objects imaged by CT are usually opaque to the naked eye, the penetrating power of the high frequency electromagnetic radiation typically used renders these objects effectively transparent. Pixel intensities obtained from CT, suitably transformed to account for attenuation, can be interpreted as proportional to the mass along the corresponding rays between the source of the radiation and the detector. Tomography is defined as the reconstruction of sectional slices of an object from such image measurements taken from different orientations [60]. The applications of CT are far-reaching and diverse, including medical imaging, airport security, and non-destructive testing in manufacturing.

The tomographic imaging process can be described mathematically by a parallel projection known as the Radon transform (Figure 4.1). The mass density $f(x, y)$ over a 2D slice of an object is projected in many different viewing directions θ giving rise to 1D images parameterized by d, as follows:

$$P(\theta, d) = \iint f(x, y)\delta(x\cos\theta + y\sin\theta - d)\,\mathrm{d}x\mathrm{d}y, \tag{4.1}$$

where $\delta(t)$ is the Dirac delta function.

To give a concrete example, X-ray CT has found wide use in medical applications due to its penetrating power through human tissue and its high contrast. Using tomographic techniques, the internal structure of the human body can be reliably imaged in three dimensions for diagnostic purposes.

There are, however, a few limitations to the technique. Good reconstructions require a great deal of data (meaning many rays from many directions), so a large detector typically needs to be spun halfway around the patient for full coverage. Quite aside from concerns of efficiency, exposure to X-rays should be limited for health reasons. Moreover, because the reconstruction is sensitive to noise, the patient is instructed to remain immobile

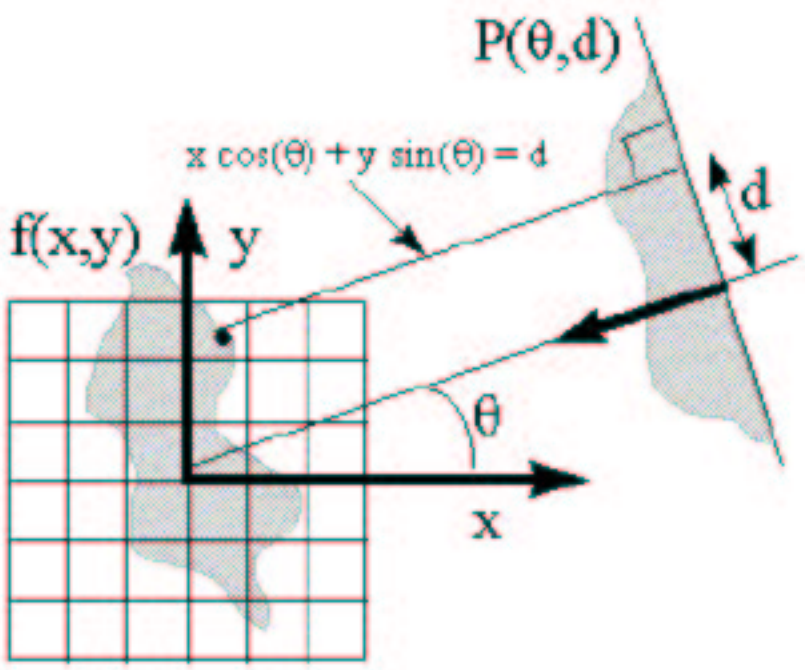


Figure 4.1: Geometry of the Radon transform.

throughout the procedure. This may be especially difficult for those injured patients for whom the CT scan is most valuable. Finally, any objects embedded in the body that are opaque to X-rays (for example, metallic shrapnel) can cause significant shadowing artefacts in the reconstruction.

4.1.1 Filtered Backprojection

The Fourier slice theorem gives a simple relationship in the Fourier domain between an object and its projections. Specifically, the theorem says that the 1D Fourier transform of a projection is equivalent to a (linear) slice of the 2D Fourier transform of the object. Moreover, this slice will pass through the origin and lie in a direction perpendicular to the direction of projection [60].

In the case of continuous images and an unlimited number views, the Fourier slice theorem can, in theory, be applied directly to obtain a perfect reconstruction. Each projection can be backprojected onto the 2D Fourier domain by means of the Fourier slice theorem, and the original object can then be recovered simply by taking the inverse 2D Fourier transform.

In practice, the CT reconstruction problem is complicated by the finite number of samples and the method's sensitivity to noise. The sensitivity to noise is due to the fact that the Radon transform is a smoothing transformation, so taking its inverse will have the effect of amplifying noise. To partially remedy this problem, the CT images are usually filtered (using some form of high-pass filter) before undertaking backprojection. These two steps, filtering and backprojection, are the essence of the filtered backprojection (FBP) algorithm which has dominated CT reconstruction algorithms for the past thirty years. FBP produces very accurate estimates of mass density for medical imaging, but many views (several hundred) are required [60] and the results are still rather sensitive to noise. See Figure 4.2 for an illustration of the FBP technique.

A noteworthy special case of FBP occurs when the object to be imaged is symmetric about an axis. Even a violently turbulent jet of flame, for example, will appear axis-symmetric if the image is averaged over a long enough period of time, provided external forces such as wind are not biased in any particular direction. Owing to the axial symmetry, the mass density $f(x, y)$ depends only on the radial distance r and so can be rewritten as $f(r)$. The 1D projection of the slice is the same from every viewing angle, and can thus be parameterized by a single variable d. This situation is described by the Abel transform:

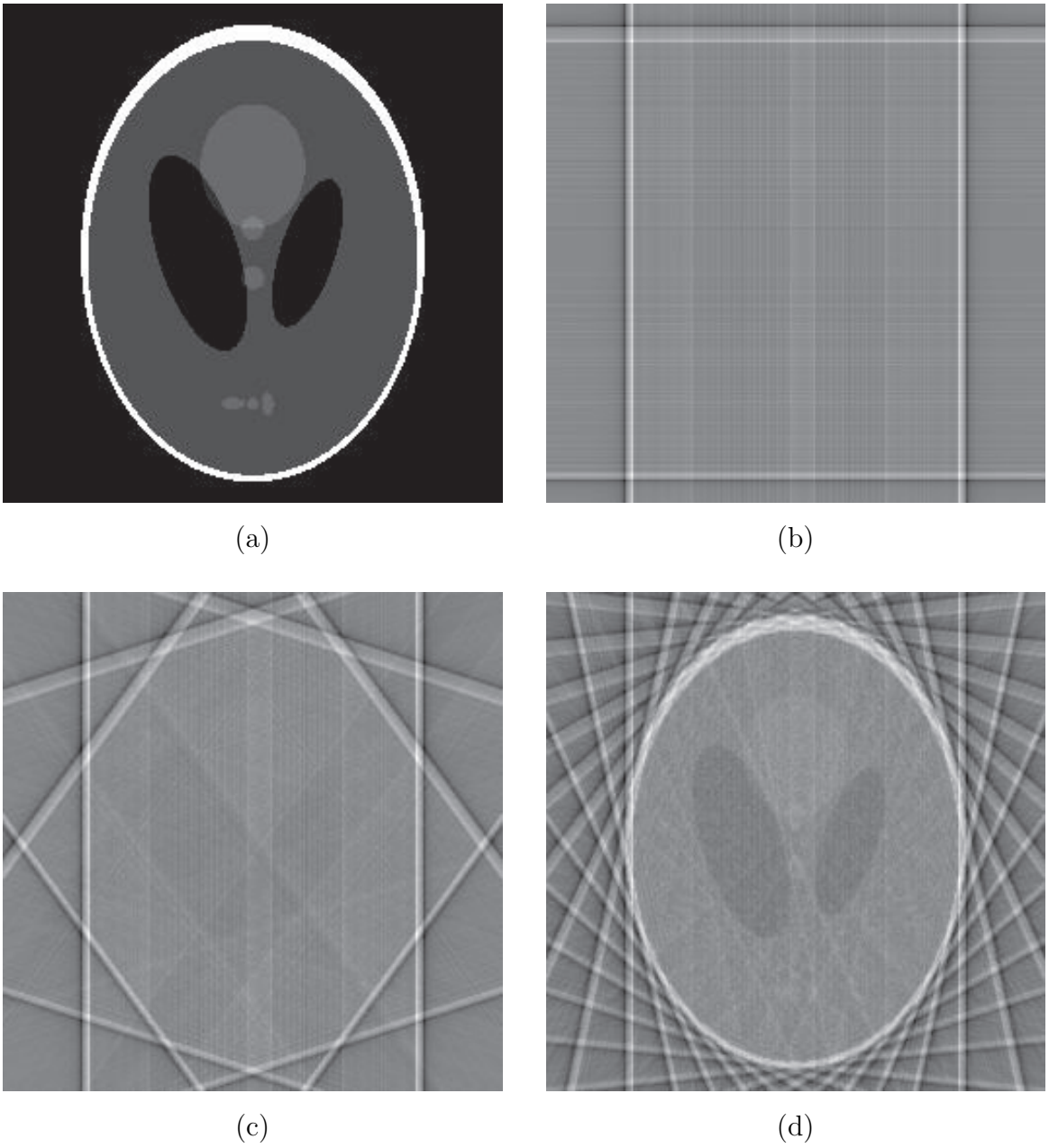
$$P(d) = 2 \int_d^\infty \frac{r f(r)}{\sqrt{r^2 - d^2}} \, \mathrm{d}r. \tag{4.2}$$

The Abel transform is straightforward to invert numerically [116], and in contrast to standard FBP, only a single projection of the object is needed to perform the reconstruction. Note that reconstructions obtained using the inverse Abel transform are especially prone to error in the central region where r is close to zero.

The FBP method has also been applied to visible light images of opaque objects with limited success [47]. This is because the image formation model assumed by the method is inconsistent with non-linear effects due to occlusion. As a result, only when the complex effects of occlusion can be ignored (for example, when the scene is convex) does the method provide reasonable approximations to 3D shape for Lambertian scenes.

4.1.2 Algebraic Methods

An alternative method for CT involves reformulating the problem in an algebraic framework. If we consider the object as being composed of an $n \times n$ grid of unknown mass densities, then each ray will impose a different linear constraint relating an image in-



(a) (b)

(c) (d)

Figure 4.2: Filtered backprojection solution for different numbers of views. (a) original image, (b)-(d) filtered backprojection with 2, 5, and 15 views. The image is a modified version of the widely used Shepp-Logan phantom [60].

tensity with a weighted sum of grid elements. Using bilinear interpolation, each input pixel depends on at most $2\sqrt{2}n$ grid elements. Reconstruction is then reduced to the conceptually simple task of finding the grid of densities which best fits the projection data, given as a collection of n-pixel 1D input images.

In the algebraic framework, the best fit is typically found through an iterative opti-

mization technique [50, 5, 60], often with additional regularization constraints to reduce non-smoothness artefacts [51, 20, 99]. Any regularization, however, comes at the cost of losing resolution in the final image. While algebraic methods are inefficient and lack the accuracy of FBP, they are far more capable of handling noisy or sparse data [60]. Algebraic methods have also been useful for handling more complicated detector geometries that involve non-parallel or curved rays [60]. These geometries are more difficult to model using basic Fourier theory.

The first method developed using the algebraic framework was the algebraic reconstruction technique (ART) [50]. Starting from some initial guess for the grid of mass densities, this method considers each of the ray constraints in turn. The difference between the measured and the computed sum along a given ray is then used to update the grid solution by a straightforward backprojection of some portion of the residual along the ray.

Unfortunately, the results obtained using basic ART are rather poor [60]. The reconstruction is plagued with salt-and-pepper noise, the convergence is slow, and the reconstruction may even be negative in places (which is physically impossible). Results can be improved somewhat by adding a relaxation parameter as in simulated annealing (see §4.1.3) at the further expense of convergence. Improvements have also been demonstrated by ordering the constraints so that the angle between successive rays is large [5], and by modifying the correction terms using heuristic error functionals that emphasize the central portions of the object (for example, adding a longitudinal Hamming window [60]).

Variations on ART, including the simultaneous iterative reconstruction technique (SIRT) [60] and the simultaneous algebraic reconstruction technique (SART) [5], differ only in how corrections from the various constraints are bundled and applied. In SIRT, all constraints are considered before the solution is updated with the average of their corrections. SART can be understood as a middle ground between ART and SIRT. It involves applying a bundle of constraints from one view at a time.

4.1.3 Statistical Methods

Statistical methods for CT are often built around the same framework as algebraic methods, however they approach the reconstruction problem in a fundamentally probabilistic manner, incorporating knowledge about the prior probability of a given mass distribu-

tion. For example, priors have been formulated to favour local smoothness [51, 99] and the presence of generalized cylinders [20]. Like the algebraic methods, statistical methods are mainly geared towards reconstruction problems where the data is noisy or somewhat sparse.

The obvious goal of imposing such a prior is to encourage the reconstruction to satisfy properties thought to be present in the real mass distributions likely to have produced the projection data. However, imposing a prior has the additional benefit of regularizing the problem. Under the algebraic framework, for any reasonable grid resolution, there are many more unknowns than equations. This is especially true when the projection data is sparse, so we are often faced with a severely ill-posed problem [51, 20, 99, 75]. While the space of mass distributions projecting to the data may be vast, the prior provides a new metric for evaluating the relative goodness of these solutions and navigating within this space.

The statistical methods differ mainly in their choice of priors and how they set about finding the maximum *a posteriori* (MAP) reconstruction. One generic technique for finding the MAP reconstruction is to use simulated annealing [75, 119, 49], a stochastic search method for high-dimensional spaces that operates by slowly decreasing the amount of randomness until the search reduces to greedy hill-climbing. While simulated annealing has guaranteed global convergence and requires no information about the gradient of the function being optimized, in practice the "cooling schedule" required to obtain good results can be very inefficient [117]. Another generic technique useful for finding the MAP reconstruction is the expectation maximization (EM) algorithm, however for this to be applicable the prior must be of a special form [88].

Sauer and Bouman presented an interesting technique called local tomography that finds the MAP reconstruction by iteratively optimizing each element in the grid of densities [99]. In order to make the problem tractable, the density is discretized to a fixed number of values (on the order of twenty). A prior is also introduced to favour local smoothness, according to some statistical description of regularity. This technique was recently extended to multiresolution for increased efficiency and robustness [46].

Wavelets have also been used to extend FBP for CT reconstruction [97, 12, 92, 127] . The basic idea is to apply the wavelet transform at the level of the 1D projections, which in turn induces a multiscale decomposition of the 2D object. For a fixed level of detail, this method is equivalent to FBP, but has the advantage of obtaining multiresolution information with little additional cost. More importantly, the wavelet method also pro-

vides a better framework for coping with noise. Bhatia and Karl suggested an efficient approach to estimating the MAP reconstruction using wavelets [12], in contrast to other more computationally intensive regularization approaches [99, 88, 75].

4.1.4 Application to Fire and Gas Reconstruction

The use of tomography for fire and gas diagnostics has been growing steadily since it was first introduced in the optics literature by Faris and Byer [41, 42]. Tomography has the advantage of being non-intrusive, but current methods often rely on complex detector configurations that limit their applicability to the optics research lab. The central problem facing traditional tomography methods for fire reconstruction remains acquiring data quickly enough to generate reconstructions with high spatial and temporal resolution. Various engineering challenges restrict the rate of data acquisition, and current reconstruction algorithms give poor results for a very sparse number of views. The number of synchronized cameras available is practically limited, and the acquisition is slowed even further if a full infrared spectrum is required for chemical species identification. While some efforts have been undertaken to reconstruct fire, to our knowledge none of these techniques have done so for the purpose of rendering and animation.

One common (but undesirable) solution to the limited data acquisition rate is to take long enough exposures, on the order of minutes, so that even the most highly turbulent flame becomes axis-symmetric when averaged over time [31, 125, 4]. Reconstruction then becomes a simple matter of inverting the Abel transform equation (4.2) for a single projection image. Unfortunately, the results for the axis-symmetric case lack the geometric detail demanded by our goal of reconstructing detailed visual models of fire.

Many of the fire reconstruction methods based on tomography employ the simple soot volume (Beer-Lambert) image formation model (§2.3.1), in which both emission and absorbance are derived from a single density field [31, 8, 124]. The necessary conditions required for the soot volume model to hold have been demonstrated empirically. Namely, the optical thickness of fire is both relatively small [31] and relatively insensitive to temperature variation [8]. The soot volume model has also been shown to agree empirically with thermocouple measurements and temperatures from two-wavelength pyrometry [31].

Baum, McGratten, and Nyden used infrared spectrometers to estimate chemical species concentration, and noted that increased smokiness caused the absorbance baseline to slant upwards with wavelength [8]. Because the spectrometers take approximately one

minute to complete a scan, the resulting temporal resolution is quite low. They then hallucinated novel views by interpolating a sparse set of actual projections, and performed regular FBP on the smoothed data. According to their analysis, based on the Nyquist theory of regular sampling, approximately three orders of magnitude speedup would be required for their method to be used for validating the Large Eddy Simulation Model of fire developed by the National Institute of Standards and Technology (NIST).

Other tomographic techniques for fire reconstruction measure deflection angle rather than emission visible to the naked eye [41, 42, 3, 4, 125]. Rainbow schlieren techniques [3] or Moiré fringe interferometry [125], for example, can be used to measure beam deflection. Deflection measurements in turn allow the refractive index field to be reconstructed using a modified version of FBP that no longer involves simple straight-line integrals [41, 3, 125]. This allows the temperature field to be derived by purely optical means by applying the Gladstone-Dale equation (2.4) to the reconstructed refractive index field [3]. In all experiments, deflections through fire were uniformly small, with no deflection higher than 0.1° observed. This is important because the modified versions of FBP assume small deflection angles to simplify the calculations.

Schwarz presented the most ambitious method for the tomographic reconstruction of fire to date [101]. His goal was the reconstruction of extinction, emission, and refractive index as three independent fields. To this end, he constructed an elaborate apparatus consisting of twenty cameras arranged in a circle, a central rotating mirror to send laser pulses in all directions, and electronically controlled schlieren filters. By taking different images with and without the fire, the laser and the schlieren filters, Schwarz was able to combine images to isolate the three fields and reconstruct them independently using tomography. The apparatus is complex and difficult to calibrate, but his results were intriguing for the complexity of the image formation model and their relatively high spatial resolution. Of course, not all scans can be taken simultaneously, so most of his experiments involved non-luminous or quasi-static flames.

In the atmospheric sciences, even less data is typically available, perhaps only a few line integrals from a limited number of Fourier transform infrared (FTIR) spectrometers that take up to five minutes per scan. With so little data, algebraic techniques like ART converge to the projection data well but suffer badly from overfitting. One remedy is to impose smoothness by encouraging the third derivatives to be small [90, 44]. Another alternative suggested by Drescher, *et al.* is fitting smooth basis functions (several 2D Gaussians) to the density field using simulated annealing [35].

4.1.5 Segmentation using Tomography

For some tomography applications, accurately segmenting 3D shape is more important than obtaining a full reconstruction [75, 53, 38, 39]. In such cases, performing the reconstruction, examining the various isocontours, and thresholding may not be required to obtain a segmentation.

In fact, by using level set methods, the 3D shape that best fits the projection data can be evolved directly, by deforming some initial set of surfaces [38, 39]. This technique also allows the densities corresponding to each level set to be evolved simultaneously. Note that this technique is closely related to the active contour method (snakes) used in computer vision for 2D segmentation [61]. Surface evolution is analogously driven by some energy function that balances accurately projecting the shape to the input images with a prior favouring low curvature, for example [38]. For one angiogram application, this prior was modified to favour long skinny cylinders, so that the arteries being segmented would not be split into smaller disconnected pieces [38].

This segmentation technique gives good results for complicated geometries and an extremely sparse number of views. As few as three to five views are needed for a good segmentation, compared to the dozens of views required by statistical reconstruction methods designed for handling noise and limited data [99, 12, 46]. One problem with the method is that the heuristic nature of the energy function could entail laborious hand-tuning of the parameters or modification of the prior for each new problem instance. Furthermore, the simultaneous evolution of more than one or two different level sets is difficult to calculate and increasingly unstable.

4.2 Perception of Transparency

Transparency arises in everyday life from a number of different physical phenomena. These include soft shadows, dark filters in sunglasses, silk curtains, city smog, fire, and sufficiently thin smoke. When we refer to transparency in a perceptual context, we usually mean that the transparent medium itself is at least partially visible. This disqualifies from our analysis the air on a perfectly clear day or a well-polished window without reflections. Since fire is a semi-transparent medium and one of our ultimate goals is to create visualizations of fire, the issues in the perception of transparency are relevant to evaluating our results.

The perception of transparency is only loosely constrained by the laws of optics. In fact, figural unity is perhaps just as important a factor in the perception of transparency as the relationship between the brightnesses of different image regions [10, 74]. Figure 4.3 illustrates how the absence of figural unity can prevent the perception of transparency actually present in the scene. Nakayama also demonstrated experimentally that transparency is fundamentally achromatic in nature [78]. Combinations of colour which are unlikely to arise in real-world scenes still give the perception of transparency.

(a) (b) (c)

Figure 4.3: Figural unity and the perception of transparency. In the regular perception of transparency (a), we observe a small transparent square above a large dark square. However, where there is an abrupt change of shape (b), the perception of transparency can break down. The small transparent square in (c) is perceived as a painted patch, presumably because this is the cognitively simpler explanation.

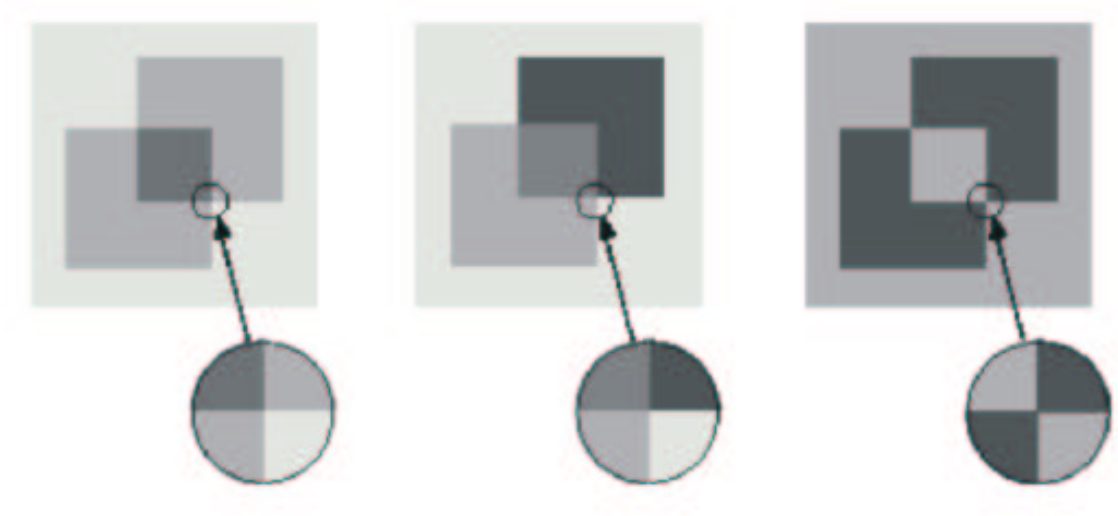
Previous theories of visual perception have described cognitive models in which more complicated images can be described economically using primitive images and combination rules. In particular, Adelson and Pentland formulated an explicit cost model for performing such a decomposition, as a concrete illustration of this idea [2]. Using this formulation, the most plausible interpretations of a scene are the cheapest interpretations (in terms of shape, lighting, and reflectance) that are consistent with the image. It seems at first glance that transparency would fit nicely into this cost-based framework, however Adelson and Anandan argue that this is not in fact the case [1]. According to them, the perception of transparency is essentially pre-physical and heuristic in nature, and not part of a full intrinsic image analysis.

4.2.1 X-Junctions

Metelli was the first to analyze constraints on the perception of transparency with layers of transparent two-dimensional shapes [74]. In his model, each layer can attenuate the luminance beneath it by a factor a, $0 < a \leq 1$, and emit its own luminance e, $e \geq 0$. Thus, the luminance at layer n is given by the relation $I_n = a_n I_{n-1} + e_n$. Note that this model is connected to the alpha channel used in computer graphics [104].

The constraints imposed by this model have been recently examined at X-junctions, places in an image where four different regions are created by the intersection of two lines. X-junctions have been shown to be especially important in establishing local constraints for scene interpretation. These local constraints may in fact propagate to constrain the interpretation of the entire scene.

It has been proposed that the human visual system employs heuristics based on X-junctions to categorize different types of transparency [1]. X-junctions can be classified into three groups based on the ordinal relationship between the brightnesses of the regions in the horizontal and vertical directions (see Figure 4.4). Satisfyingly enough, a mathematical analysis of Metelli's rule at X-junctions leads to constraints which justify heuristics based on the degree of reversingness.



(a) (b) (c)

Figure 4.4: Different types of X-junctions. The non-reversing case (a) and single-reversing case (b) are both readily perceived as transparent, but the depth ordering in the non-reversing case is ambiguous. The double-reversing case (c) is not perceived as transparent. Adapted from [1].

Note that transparency has also been shown to be eagerly perceived by untrained observers even in a number of physically impossible situations, as when the ordinal relationships required at X-junctions are violated [10]. However, part of this effect may be due to demand characteristics in the experimental design.

While some heuristics for the perception of transparency (figural unity, X-junctions) have already been investigated, their full characterization remains an open problem. The observation that children are terrified of falling through transparent floors suggests that our perception of transparency may be learned by experience [48]. One promising avenue of research may be to follow the early development of heuristics governing the perception of transparency in animals and the young.

4.2.2 Perception of Transparency in 3D

The difficulty in making judgments about 3D transparent shapes has been demonstrated experimentally [54]. One technique suggested for improving the visualization of 3D transparent shapes involves overlaying opaque strokes at different points on the surface of the shape, where the orientation and size of the strokes are chosen to correspond to the direction and magnitude of principal curvature [54]. This texturing is shown to improve performance in judging relative distances for a medical visualization task.

Interestingly, these results suggest that the reconstruction of a transparent 3D object may not need to be as accurate as that of an opaque object of similar shape to be convincing. Even for transparent objects, however, we expect poor reconstructions to be betrayed by smoothly varying the viewpoint, or if salient features such as contour or significant refractive effects are not preserved.

4.3 Matting Techniques

Matting is another important application for transparency. The process of matting involves separating foreground objects from the background, typically so that they can then be composited over a new background [105]. In matting, transparency is modelled on a per-pixel basis, rather than in scene space.

The end result of the matting process is an image (a matte) that is transparent in regions of the foreground object, opaque in the background regions, and semi-transparent at object boundaries and wherever the foreground object is semi-transparent. This is

completely analogous to the alpha channel used in computer graphics for compositing [104].

4.3.1 Blue Screen Matting and Extensions

Historically, the driving force behind matting technology has been the film industry, which employs these techniques to produce special effects. Vlahos pioneered the well-known blue screen matting technique over forty years ago, and refinements of this technique are still in use today. The basic idea is to film the foreground object (usually an actor) in front of a bright screen of known colour (usually blue). If the object is relatively unsaturated, perhaps near-grey or flesh coloured, a simple equation involving the background and observed foreground colours allows the transparency to be calculated at every pixel [105]. Human intervention is usually needed to identify a good matte when using commercial matting machines.

Smith and Blinn demonstrate mathematically that the general matting problem is unsolvable, but justify traditional approaches to matting by showing how certain constraints reduce the space of solutions [105].

They also suggest a novel approach to matting that they call triangulation. This method requires the object to be filmed against two or more backgrounds which must differ at every pixel. The solution to the matting problem then becomes overdetermined, so that fitting can be performed in the least-squares sense. The main difficulty with the method is its lack of applicability when only one camera is available and either the object or the camera is moving. If the object is static and the camera is computer-controlled, then in theory two separate passes can be made, however maintaining proper calibration remains a issue.

There is a whole new class of matting techniques that use a Bayesian framework to estimate the maximum *a posteriori* (MAP) decomposition of each pixel into a mixture of foreground and background. These techniques have had success computing alpha mattes using one image and a modest amount of user interaction [26, 28], or by using multiple calibrated cameras [122]. Motion information over a video sequence can also be used to further constrain the matting [28, 122].

4.3.2 Environment Matting

The environment matting framework proposed by Zongker *et al.* extends these ideas even further in order to capture the refractive and reflective properties of foreground objects [128]. The impressive results obtained using this method (for static objects) are practically indistinguishable from real photographs. While additional research has been done in capturing environment mattes from moving objects in front of a single known background, the simplifying assumptions required to make this possible cause the accuracy to suffer significantly [27].

The basic environment matting method operates by photographing the foreground object with a series of structured images behind it and to the sides of it. Then, for each pixel, a non-linear optimization problem is solved, which attempts to find the coverage of the foreground object and the axis-aligned rectangle in the background which best reproduces the pixel colour over all of the input images.

The backgrounds are chosen to be a hierarchical set of horizontal and vertical striped patterns, corresponding to one-dimensional Gray codes. This simplifies the calculation of average value within axis-aligned rectangles in the background, and reduces the dimensionality of the optimization to a more manageable three.

In further work, additional accuracy was obtained by using swept Gaussian stripes as background images [27]. By considering stripes with different orientations, the environment matte could be recovered as a set of oriented elliptical Gaussians instead of axis-aligned rectangles. This allows much greater realism for objects made from anisotropic materials, such as brushed steel.

4.4 Transparency in Shape Recovery and Motion Analysis

A variety of computer vision algorithms, such as estimating depth from stereo, recovering 3D shape, or computing optical flow, have recently been extended to handle some degree of transparency [113, 7, 55, 112, 58]. Most of these extensions, however, were not motivated by the desire to handle transparent objects in general. Rather, the typical goal was to improve behaviour at troublesome mixed pixels, which are the result of integrating radiance over surfaces with radiance discontinuities or across an occlusion boundary. Mixed pixels are also caused whenever transparent objects are present in the scene.

Properly coping with mixed pixels involves managing multiple explanations per pixel, so these methods tend to be more complicated than their predecessors. Transparency is typically modelled by these methods as alpha mattes (see §4.3) for a small number of layers making up the scene. As such, their performance with more general transparent objects requiring a full volumetric description of transparency, such as smoke or fire, is limited.

4.4.1 Traditional Stereo with Mixed Pixels

Traditional stereo methods have recently become more robust to the presence of mixed pixels, especially those at occlusion boundaries due to limited sampling. Birchfield and Tomasi addressed the problem of mixed pixels with an efficient metric for measuring the similarity between corresponding pixels in a stereo pair [14]. Their metric is provably insensitive to sampling and thus better handles mixed pixels due to limited resolution.

Szeliski and Golland presented a stereo algorithm which operates over discretized scene space, but also adds an explicit representation for partially transparent regions [113]. Like other volumetric stereo methods, evidence for competing correspondences is considered. This method differs in allowing the detection of mixed pixels at occlusion boundaries in order to obtain sub-pixel accuracy. Estimates for colour and transparency are refined using a global optimization method which penalizes non-smoothness and encourages transparencies to be either 0 or 1.

Another stereo algorithm which accounts for transparency was proposed by Baker, Szeliski, and Anandan [7]. Their method models the scene as a collection of approximately planar layers, or sprites. These sprites are initially estimated without considering transparency, then this solution is refined iteratively. First, residual depths perpendicular to the planes are estimated, then per-pixel transparencies are estimated as well. The images are re-synthesized using a generative model that incorporates transparency, and error with respect to the input images is minimized using gradient descent.

These new stereo methods achieve gains over previous results, and this improvement may be directly attributed to their better handling of mixed pixels. However, none of these methods present results for scenes containing real transparent objects.

4.4.2 Volumetric Reconstruction with Uncertainty

The voxel colouring algorithm of Seitz and Dyer [102] was inspired by early work on volumetric reconstruction from images (of opaque objects) that properly accounted for visibility [66, 30]. Kutulakos and Seitz later extended this method to obtain the space carving algorithm [64]. Both algorithms operate on the principle that a set of cameras must agree on the colour of an opaque voxel, though only if that voxel is visible from all of those cameras. In this framework, voxels are labelled as either completely transparent or completely opaque. Note that even for opaque objects, a perfect voxel reconstruction would require transparency (partial occupancy) information to be computed for the voxels at the boundaries of scene objects.

Accounting for mixed pixels is important for volumetric reconstruction as well. The space carving technique assesses visibility by determining whether a given voxel projects to roughly the same colour in a set of images [64]. However, if a given voxel is located at a depth discontinuity with different backgrounds from different views, then the resulting (mixed) pixel projections may appear inconsistent and lead the algorithm to label the voxel (incorrectly) as empty. The algorithm may similarly fail in the presence of other sources of uncertainty such as poor camera calibration or sophisticated lighting effects not captured by the image formation model.

The approximate space carving method presented by Kutulakos tried to address this problem by broadening the notion of colour consistency between image pixels [63]. He described the shuffle transform, a looser matching criterion where arbitrary reordering of pixels within a local neighbourhood is permitted for the evaluation colour consistency. Other approximate matching criteria such those based on rank order within a local neighbourhood [126] may also improve performance in this respect.

Space carving has lately been extended using probabilistic methods to permit the representation of partial or expected voxel occupancy [13, 21]. The basic space carving algorithm was modified to incorporate stochastic sampling, and the algorithm is shown to generate fair samples from a distribution which accurately models the probabilistic dependencies between voxel visibilities. This approach provides a principled manner of dealing with the uncertainty in the volumetric reconstruction, although mixed pixels are simply modelled as being due to noise.

4.4.3 Volumetric Reconstruction with Transparency

Several recent volumetric reconstruction techniques have explicitly attempted to recover transparencies along with voxel colouring information [56, 15]. In these techniques, image pixels are generated from weighted combinations of voxel intensities along corresponding rays, where the weights are a function of the voxel transparencies.

The major shortcoming of these techniques is that the reconstruction methods are not entirely consistent with the imaging model being used. Notably, voxel visibility is assessed only by examining whether the voxels along a viewing ray are semi-transparent. This ignores long-range global dependencies between voxel visibilities [13], leading to algorithms whose results are difficult to characterize in terms of the physical properties (shape and transparency) of the underlying scene. These methods are also restricted in their ability to handle very noisy or sparse projection data.

Dachille *et al.* reported one such reconstruction method incorporating transparency that they call volumetric backprojection [56]. The method is a straightforward extension of SART (see §4.1.2) made efficient with hardware acceleration and a parallel architecture for performing backprojection on a slice-by-slice basis. If each image is available with both black and white backgrounds, a simple matting technique (see §4.3) can be applied to extract the cumulative effect of transparency on each pixel. Then, correction terms for both voxel colours as well as voxel transparencies can be backprojected onto the reconstruction volume. However, the physical basis for this backprojection step is unclear, especially because the scene can contain opaque voxels and the long-range visibility dependencies between them are ignored.

Results for the volumetric backprojection algorithm were demonstrated for synthetic data, where the exact lighting conditions are made available to the reconstruction algorithm. While this technique represents an interesting proof of concept, it is unclear that this technique would perform acceptably on real reconstruction tasks, where noise and unknown lighting conditions may be additional complicating factors.

Another volumetric reconstruction technique incorporating transparency is the roxel algorithm proposed by De Bonet and Viola [15]. In this approach, responsibilities (weights) are assigned along rays to describe the relative effects of the voxels in determining observed pixel colour. The responsibilities are thus due to the cumulative effect of transparency in attenuating the colours of the voxels. While the relationship between responsibilities and transparencies is nonlinear, a closed form is available for

converting directly between the two representations.

In the roxel algorithm, the entire volume is initially set to be empty and completely transparent. Each iteration of the method comprises the following sequence of three steps. First, colours are estimated using a generalization of backprojection, in which voxel colours are weighted by their responsibilities. Then, responsibilities are estimated based on the disagreement between the estimated pixel colour and the real image data. To distribute responsibility among the voxels in each ray, a softmax style assignment (as in reinforcement learning) is made. Voxels that agree better with the data are given a greater share of the responsibility, and the overall distribution is controlled by a single parameter expressing belief in the noisiness of the data. Finally, both the responsibilities and the related transparencies are renormalized to be globally consistent over the entire set of view estimates.

Results with the roxel algorithm were demonstrated for a variety of synthetic and real data, however the reconstructions that involved transparent objects were not as accurate. Another interesting (but physically unfounded) aspect of the roxel algorithm is that uncertainty in the reconstruction is also represented by transparency. Thus, an opaque but uncertain voxel will be rendered as semi-transparent, corresponding to its expected value. This leads to an algorithm whose results are difficult to characterize because it confounds various physical properties of the underlying scene.

4.4.4 Motion Analysis with Transparency

Motion analysis algorithms have lately been extended to handle the motions of transparent objects. Irani, Rousso, and Peleg presented a recursive high-level motion analysis technique, geared at extracting a small number of dominant motions, where the objects may be occluded or transparent [55]. First, the dominant affine motion is estimated using temporal integration, producing an integrated image that is sharp for pixels involved in the dominant motion and blurred for the others. The sharp pixels are then segmented and removed from the original sequence, so that the algorithm can estimate the next most dominant motion.

The temporal integration technique is robust to occlusion, and the algorithm can even handle certain types of transparent motions. The success of the algorithm in the presence of transparency depends on different levels of contrast to allow the identification one dominant transparent motion. Success also depends on the assumption that the dominant

transparent motion can be approximately removed from the sequence by considering the difference between the original frames and the integrated image. Good results are presented for a picture frame partially reflecting objects behind the camera.

More recently, Szeliski, Avidan, and Anandan also addressed the extraction of component layers from a sequence of images containing reflections and transparency [112]. Their method likewise estimates a small number of global parametric motions describing how the layers move in relation to each other. However, unlike the temporal integration technique, which leaves the component layers blurred together, their method solves a constrained optimization problem and obtains a full separation of the different components.

The skin and bones method suggested by Ju, Black, and Jepson also handles transparency in its estimation of optical flow [58]. Robust statistical techniques and mixture models are used to estimate multiple layers of motion within fixed patches of the scene, and regularization constraints ensure that smoothness between nearby and related patches is maintained.

Transparency is handled by the skin and bones method in that multiple depth measurements are allowed to exist for each pixel, and the regularization step will automatically segment the data into different surfaces. The skin and bones method can be viewed as a hybrid between dense optical flow methods and global parametric models which extract only the dominant motions.

4.5 Vision in Bad Weather

In computer graphics, sophisticated physically-based models have been suggested to animate and render atmospheric phenomena such as clouds, fog, and smoke [98, 107, 45, 43]. One proposed method for visualizing smoke was based on radiosity style techniques to approximate the effects of multiple scattering [43]. Computer vision techniques, on the other hand, often assume that viewing conditions are clear and are rarely explicit in considering the effects of atmospheric phenomena on image formation. Similarly, most image formation models of fire (see §2.3) assume smokeless conditions or consider the effect of smoke separately.

Nayar and Narasimhan have proposed that the effects of bad weather should be considered when designing vision systems [80, 79]. Moreover, they describe how bad weather conditions might even be turned to an advantage, if we take the view that the

atmosphere acts to modulate information about the scene. They describe algorithms that estimate additional information about the scene, such as relative depth, given an atmospheric model.

This extraction of additional information is made possible by making simplifying assumptions about the atmospheric model. Specifically, these techniques work best under uniformly hazy conditions, and do not cope well with more sophisticated scattering phenomena. For example, the effect of sunlight selectively breaking through certain areas of low-lying clouds would be modelled quite poorly using these techniques. In the same way, we would expect the reconstruction of fire to be relatively unaffected by a uniform blanket of thin smoke, but more upset by thicker and heterogeneous smoke.

Using multiple images of the same scene taken under different weather conditions, relative depths of point light sources can be estimated by comparing the degree of attenuation in the images [80]. If the images are taken at night, illumination from the environment is negligible and a simple scattering model suffices. The optical thicknesses of the different weather conditions can then be easily computed in order to estimate the relative depths.

Another technique involves estimating relative depths from a single image of a scene that is blanketed in dense grey haze [80]. In this situation, the dominant source of image irradiance is a phenomenon known as airlight, corresponding to inscatter in the radiative transfer equation (2.8) where the albedo is significant. Airlight causes the hazy atmosphere to behave as a source of light through the scattering effect of its constituent particles, so that brightness will increase as a function of path length. By measuring the image intensity at the horizon, the airlight for an infinite path length can be approximated. The simple airlight model can then be used to estimate the relative depths corresponding to different image intensities.

Narasimhan and Nayar also derived a dichromatic model for atmospheric scattering, showing that the colour of a point under bad weather conditions is a linear combination of its clear-weather colour and its airlight colour [79]. Because the colour-space direction of the airlight is constant over the entire scene, it can be computed robustly by intersecting the dichromatic planes corresponding to the different colours appearing in the image. This also allows the relative magnitudes of airlight colours to be determined for all the points in the scene. Then, given multiple images of the scene taken under different weather conditions, relative depths can be estimated even more reliably by using only the airlight components of the image colours. Moreover, true colours can be determined

for the entire scene given the clear-weather colour of just a single point in the image.

4.6 Summary

For the purposes of visual reconstruction, fire should be understood as a luminous transparent medium. To this end, we have examined transparency as it has been addressed by various areas in computer vision. Tomographic reconstruction is directly relevant because of the similar imaging model involved, though the severely limited number of views modifies the nature of the fire reconstruction problem somewhat. Other computer vision techniques involving transparency are less relevant, either because the representation of transparency used is not volumetric or because certain assumptions that are made are inconsistent with the physical model of fire. The research on perception of transparency is interesting and may later be used to help evaluate the quality of different fire reconstructions.

While transparency is a fertile area for research in computer vision, the accurate modelling of transparency is by no means the end goal. Research on transparency is being broadened to even richer models of light transport, including such phenomena as refraction, reflection, and scattering. The work on vision in bad weather and the environment matting technique are early examples of this trend. As the modelling and recovery of physical scene attributes continues to improve, the gap between the representations used in computer vision and those used in computer graphics will likely continue to narrow over the next decade.

Chapter 5

Image-Based Reconstruction and Rendering of Fire

> The shape of fire is a dance;
> the woman is heat lightning
> whose summer dancing shapes the fire
> into the dancer's hand, a shaking wrist,
> and limbs that wave like forest mosses
> in the wind.
>
> *The Shape of Fire* (l. 1–6)
> John Horvath Jr.

> Three moves equals a fire.
>
> Pioneer saying about moving one's belongings

Instead of the simple temperature and deflection measurements described in §2.2, we would like to recover a complete spatio-temporal description of fire. In its most general form (see Chapter 2) this would mean extracting the temperature, density, index of refraction, and chemical composition for every 3D point, at every instant in time. Moreover, since fire is a dynamic phenomenon, a full description would also need to include a vector field of instantaneous particle flows, as discussed in Chapter 3.

Our goal is to derive this spatio-temporal description of fire from images or videos of fire taken simultaneously from different viewpoints (see Figure 5.1). This leads to an extremely complex visual recovery problem for which no current solution is known. While it is not clear how this problem in its full generality can be solved, there are a few key properties that we believe any approach attempting to solve this problem must possess:

Physically-consistent image formation model The image formation model should

be in agreement with the optics of fire.

Instantaneous fire reconstruction The reconstruction should be possible over very small (sub-second) time intervals.

High spatial resolution The reconstruction should be have high spatial resolution, in order to be useful for computer graphics.

Sparse viewpoint sampling Few synchronized cameras should be required.

Least-commitment principle Only physically-consistent assumptions should be imposed, with no *ad hoc* constraints on the shape of the fire.

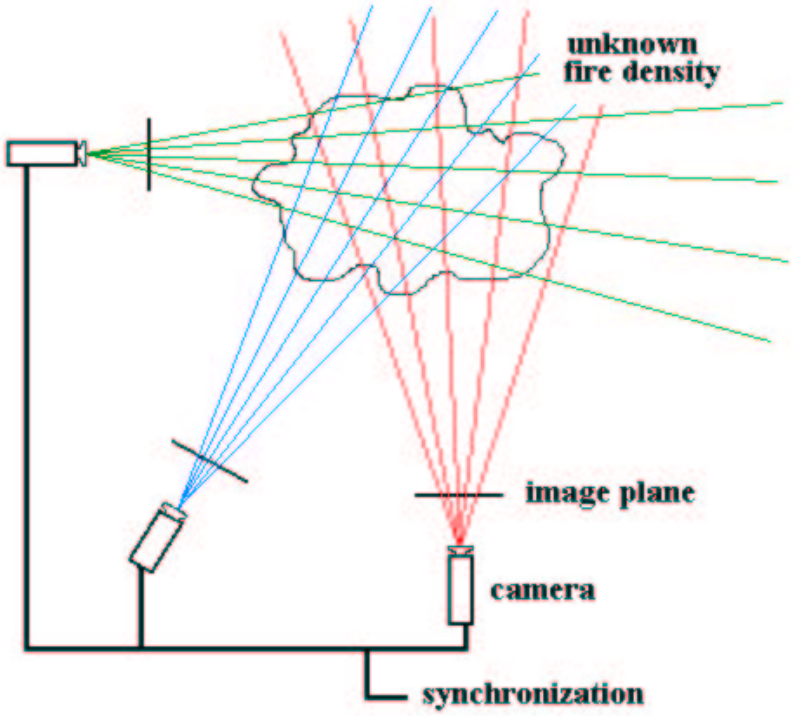


Figure 5.1: Schematic illustration of the reconstruction problem.

In this chapter, we present a new image-based method for the visual reconstruction of fire that satisfies all of these properties. Our method is visual in the sense that we do not directly resolve the fire into various physical properties such as temperature and index of refraction. However, the method is based on a physically-consistent image formation model that allows us to convincingly render the fire from novel points of view.

We show that under the assumed image formation model of fire, the reconstruction problem is equivalent to computerized tomography, but with the special difference that

many fewer views are available than in typical tomography applications. Because the cameras must be synchronized to achieve a high temporal resolution, the number of views is limited by the number of cameras available (see Figure 5.2). By contrast, many medical imaging applications use hundreds of input images.

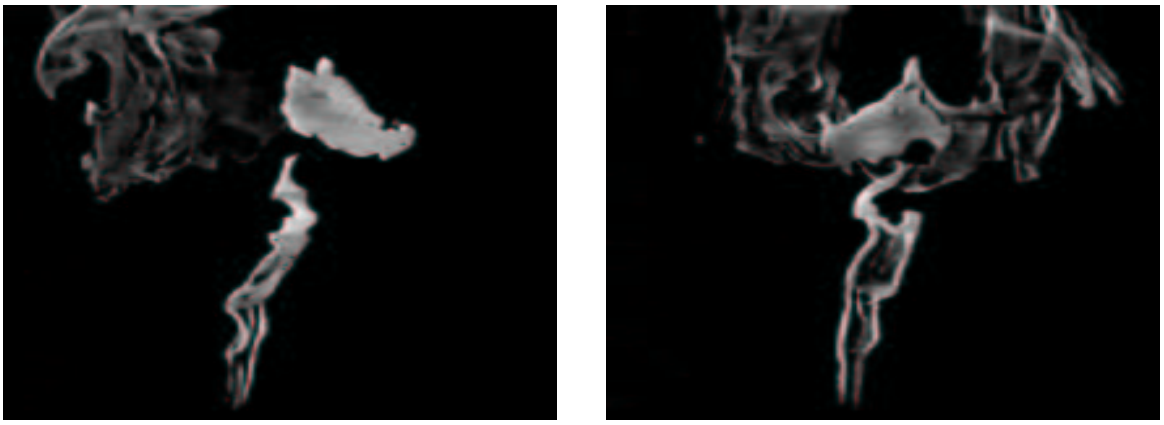


Figure 5.2: MATRIX dataset, example synchronized input images for the reconstruction method.

In order to make the problem more computationally feasible, we consider one 2D slice at a time. From the epipolar geometry between two cameras, slices can be chosen so as to project to corresponding lines in the images. For more than two cameras, a given slice may project to some larger region of an input image. With a parallel projection camera model and cameras arranged with cylindrical geometry, slices will simply project to the scanlines of the images.

The key ideas of the reconstruction method are as follows. In order to reduce the dimensionality of the problem and enforce spatial coherence, we use blobs to represent the reconstructed fire. We then demonstrate that by fixing certain blob parameters, the reconstruction is reduced to solving a constrained linear least-squares problem. Finally, we describe a stochastic resampling method for modifying the previously fixed blob parameters based on previous reconstruction results. The new reconstruction obtained using the resampled blobs parameters is a major improvement in terms of both resolution and image error.

Another contribution of this work is an analysis of the space of reconstructions that project to the input images. The only analysis of this kind that we are aware of concerns density fields whose elements are restricted to discrete values [53]. We demonstrate with continuous-valued density fields that the reconstruction of fire is fundamentally

ambiguous and show dramatic examples of this ambiguity for the two-camera case. We also present a novel image-based rendering algorithm to exploit the original input images in rendering the reconstructed fire.

Previous research relevant to fire on reconstruction from images was described in Chapter 4. While computerized tomography has already been explored for the reconstruction of fire and other gases, current methods are too low in spatial and temporal resolution [31, 124, 41, 3, 125, 35, 90, 44] or require too many cameras [101, 8] for our purposes. In the computer vision literature, a variety of techniques incorporating some model of transparency have also been proposed. However, in most cases, the model of transparency used is not rich enough to describe volumetric transparent phenomena such as fire [105, 128, 113, 7, 55, 112, 58, 80, 79]. In other cases, where transparency is indeed represented volumetrically, the proposed reconstruction methods are not entirely consistent with the imaging model [56, 15].

In order to resolve fire into various physical properties, more synchronized measurements that go beyond regular photography are needed. One realization of a more complete reconstruction of fire was Schwarz's method [101], which involved using a complicated apparatus to reconstruct the refractive index as well the density of fire (see §2.3.3 and §4.1.4). However, this method requires an impractical amount of hardware to obtain high spatial resolution, and can only perform the full reconstruction for axis-symmetric or non-luminous fire. Note that recovering a full physical model of fire also involves the non-trivial problems of inferring the sources of fuel and characterizing the turbulent wind field.

5.1 Image Formation Model

The image formation model of fire we use for the reconstruction algorithm is the simplified radiative transfer model, as discussed in §2.3.2. This model assumes negligible atmospheric scattering, which is reasonable provided that the fire is not obscured by substantial smoke. It also assumes that emission, like absorption, is directly related to the density of the fire reaction products. This coupling between emission and absorption is consistent with the soot volume model of fire as well (see §2.3.1) in which emission is the result of luminous soot particle density.

We now show how the input images can be transformed to obtain a reconstruction problem equivalent to computerized tomography. Note that only greyscale intensity

images are considered here, though colour reconstruction will be discussed in a later section (§5.3.7). Revisiting the simplified radiative transfer model of Equation (2.9), we have that the image intensity I corresponding to a given ray ℓ is:

$$I = (1 - \tau)\, Q_0, \tag{5.1}$$

where Q_0 is a constant related to the emission per unit mass. As before, the transparency τ is given by the following line integral along the ray:

$$\tau = \exp\left(-\sigma_t \int_\ell \rho(x, y)\, \mathrm{d}x\mathrm{d}y\right). \tag{5.2}$$

The transparency describes the degree to which the density field $\rho(x, y)$ both attenuates light (when viewed as an absorber), and emits light (when viewed as an emitter). The extinction cross-section σ_t is a constant relating density and transparency, and is assumed to be a uniform throughout the medium.

Next, observe that as the luminous fire mass increases without bound, the transparency τ approaches zero. From Equation (5.1) the maximum image intensity corresponding to an infinite amount of fire I_∞ is thus Q_0. We assume that our imaging system saturates at a slightly lower intensity level $I_{max} = I_\infty - \epsilon$ corresponding to some finite amount of fire. In practice, with the intensity in the range 0–255, we use $\epsilon = 1$. This saturation model lets us further simplify the imaging model for fire by eliminating the unknown Q_0:

$$I = (1 - \tau)\, I_\infty. \tag{5.3}$$

By manipulating this equation, we derive a transformation of the image intensity I' such that the transformed intensity for some particular ray ℓ corresponds to the line integral over that ray through the density field:

$$\begin{aligned} I' &\triangleq -\log(1 - I/I_\infty) \\ &= \sigma_t \int_\ell \rho(x, y)\, \mathrm{d}x\mathrm{d}y. \end{aligned} \tag{5.4}$$

In other words, we have shown that the imaging model for fire (suitably transformed) is equivalent to the imaging model used in computerized tomography (§4.1). Note that we only expect to reconstruct the density field up to a constant scale factor, namely the unknown extinction cross-section σ_t.

Up to this point we have implicitly assumed that the images of fire were obtained with completely black backgrounds. However, we can generalize the imaging model to

deal with arbitrary backgrounds, provided images of the background without the fire are available. In this case we have:

$$I = \tau I_{bg} + (1 - \tau) I_{\infty}, \tag{5.5}$$

where I_{bg} is the outgoing intensity of the background where the ray passing through the fire meets some background surface. The transformed image intensity for arbitrary backgrounds:

$$I' \triangleq -\log\left(\frac{I - I_{\infty}}{I_{bg} - I_{\infty}}\right), \tag{5.6}$$

can be derived analogously to Equation (5.4).

5.2 Analysis of Ambiguity

The reconstruction of fire (and tomography in general) is a fundamentally ambiguous problem because a large space of density fields reproduces the same input images. Figure 5.3 illustrates an elementary example of this ambiguity. While we might expect ambiguity to become gradually resolved with the addition of more input images, for the sparse viewpoint sampling assumed by our method, the consequences of this ambiguity are significant. We analyze the space of (non-negative) photo-consistent reconstructions, paying special attention to the two-camera case.

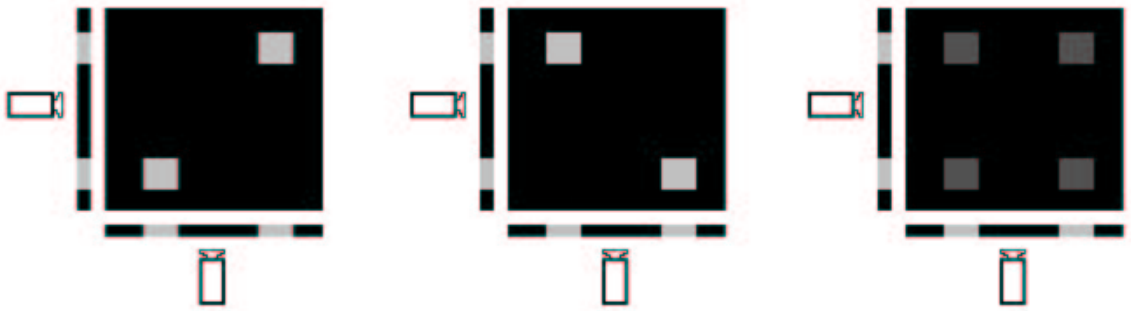


Figure 5.3: Ambiguity in the two-camera reconstruction problem. The dark grey is half the intensity of the light grey, and all three density fields shown reproduce the same input images.

We consider the 2D tomography problem, where the goal is to reconstruct an $n \times n$ field of densities projecting to a small set of n-pixel 1D images taken from different orientations. As in the algebraic framework for computerized tomography (§4.1.2), each image

pixel can be related to a weighted sum of $O(n)$ elements lying close to the corresponding line through the density field.

For clarity of presentation, we normalize all 1D input images to sum to one. Note that image sums are invariant to camera orientation because they correspond to the sum of all elements in the 2D density field. Namely, for any two (column vector) images X and Y, we have that $\mathbf{1}^T X = \mathbf{1}^T Y$ where $\mathbf{1} = [1, 1, \ldots, 1]^T$.

5.2.1 Space of Two-Camera Reconstructions

With just two input images the problem of ambiguity is especially severe. Without loss of generality, assume that the input images represent two orthogonal views aligned with the coordinate system of the density field and that we use a parallel projection camera model. In this case, the images simply correspond to the row and column sums of the density matrix. Note that for two general input views, the problem can be reduced to the orthogonal and orthographic case by a known and invertible 2D warp [123].

Now given any particular photo-consistent reconstruction S, the space of all two-camera reconstructions can be represented as $S + D$ where D is a matrix satisfying the following two properties:

Row and column sums are preserved We require that $\mathbf{1}^T D = 0$ and $D\mathbf{1} = 0$, so that $S + D$ still projects to the input images.

Non-negativity is preserved We require that $D(i, j) \geq -S(i, j)$ for all matrix elements (i, j), so that $S + D$ is still physically meaningful density function.

Note that fixing the upper-left block of $(n - 1)^2$ elements of D will uniquely define the remaining $2n - 1$ elements so that the rows and columns sum to zero.

The ambiguity associated with two-camera reconstructions has been previously described by elementary transformations known as *switching operations* [53], and we extend this analysis to continuous valued densities. A switching operation is defined to be the addition of some switching matrix Δ to the current solution, where Δ has four non-zero

entries in a rectangle aligned with the grid as follows:

$$\Delta = \begin{pmatrix} & \vdots & & \vdots & \\ \cdots & +\alpha & \cdots & -\alpha & \cdots \\ & \vdots & & \vdots & \\ \cdots & -\alpha & \cdots & +\alpha & \cdots \\ & \vdots & & \vdots & \end{pmatrix}. \tag{5.7}$$

This has the effect of shifting mass from two diagonal corners of the rectangle to the other two corners. However, we still must mind the restriction that applying a switching operation should not make the density field negative anywhere.

Theorem 5.2.1 *The space of photo-consistent two-camera reconstructions belongs to an $(n-1)^2$-dimensional hyperplane lying in the space of $n \times n$ real matrices $M_{n\times n}(\Re)$. This hyperplane is spanned by a canonical basis of $(n-1)^2$ switching matrices, and is restricted to the region where all coordinates of $M_{n\times n}(\Re)$ are non-negative.*

Proof See Appendix A for details.

Corollary 5.2.2 *Any two solutions to the two-camera reconstruction problem S_1 and S_2 can be related by a set of $(n-1)^2$ switching operations that transforms S_1 into S_2, namely $S_2 = S_1 + (\Delta_1 + \Delta_2 + \cdots + \Delta_{(n-1)^2})$.*

5.2.2 Multiplication and Sheet Solutions

We now develop two specific solutions to the two-camera reconstruction problem that represent extremes of the space of ambiguous reconstructions. As depicted in Figure 5.4, these two solutions illustrate the severity of the two-camera ambiguity problem.

Theorem 5.2.3 (Multiplication solution) *If X and Y are input images corresponding to the row and column sums of the density field respectively, then the outer product $S_{mult} = XY^T$ is a solution to the reconstruction problem. This solution is the maximum entropy and minimum rank reconstruction.*

Proof S_{mult} is a valid solution because it reprojects to the two original input images and is non-negative everywhere. Specifically, since the input images are normalized to sum to one, we have $\mathbf{1}^T S_{mult} = \mathbf{1}^T XY^T = Y^T$ and similarly $S_{mult}\mathbf{1} = XY^T\mathbf{1} = X$.

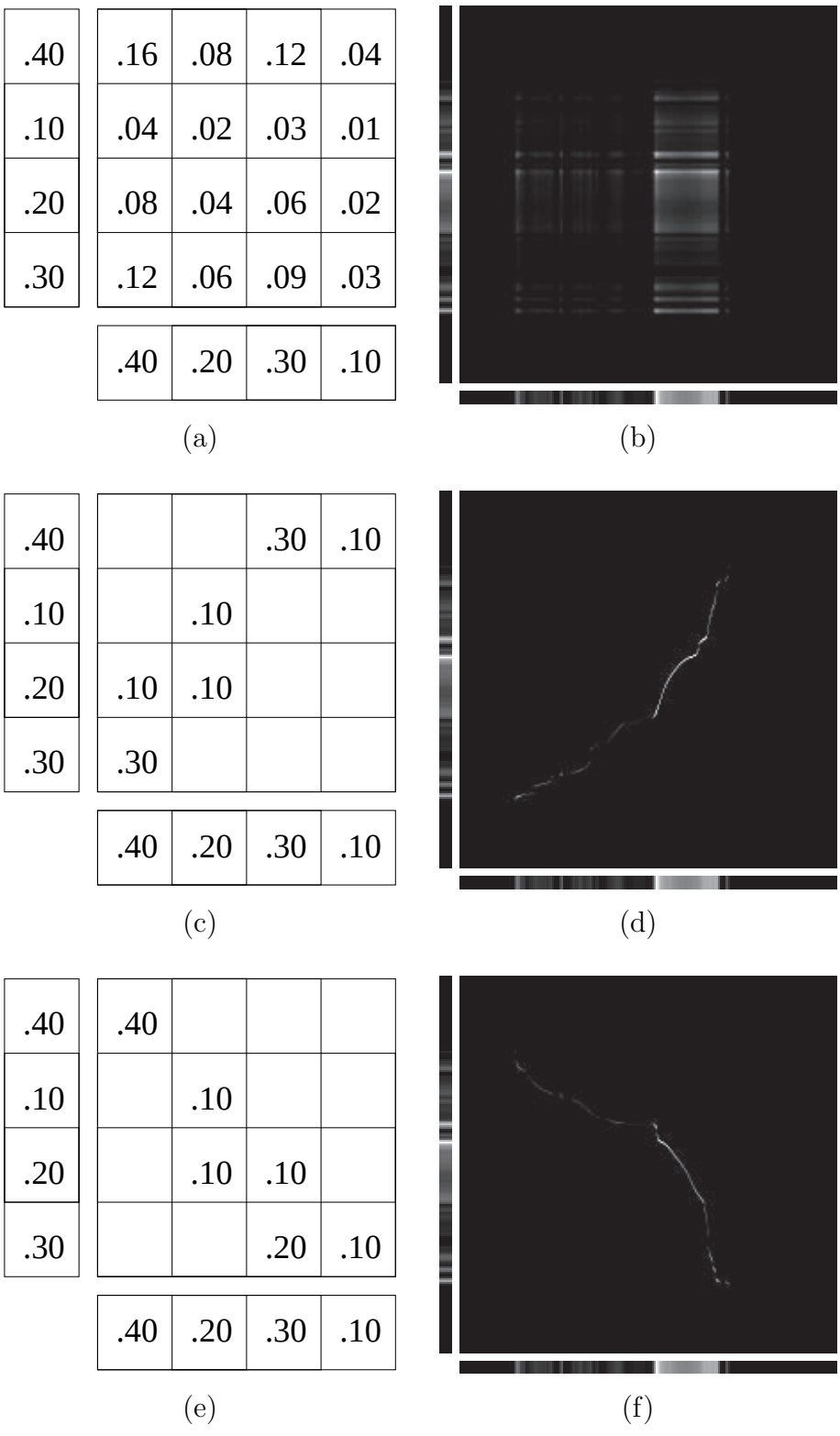


Figure 5.4: Multiplication and sheet solutions. The multiplication solution (a)-(b) and the two sheet solutions (c)-(f) are shown for a small numerical example (left) as well as a slice of the dataset shown in Figure 5.2 (right). All three solutions reproduce the same input images.

Moreover, by construction, the non-negativity of the input images ensures that S_{mult} is non-negative. Note that this solution is clearly minimum rank.

To see why this solution maximizes entropy, first observe that because the input images sum to one, we have $\sum S_{mult} = \mathbf{1}^T S_{mult} \mathbf{1} = 1$. Since S_{mult} is also non-negative everywhere, S_{mult} can be thought of as discrete probability function. From information theory, the joint entropy for two random variables X and Y is defined as $H(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) / \log P(x,y)$, where $P(x,y)$ represents the probability of a particular instantiation of the random variables. The joint entropy is maximized when X and Y are independent, namely when $P(X,Y) = P(X)P(Y)$. This corresponds directly to the definition $S_{mult} = XY^T$. ∎

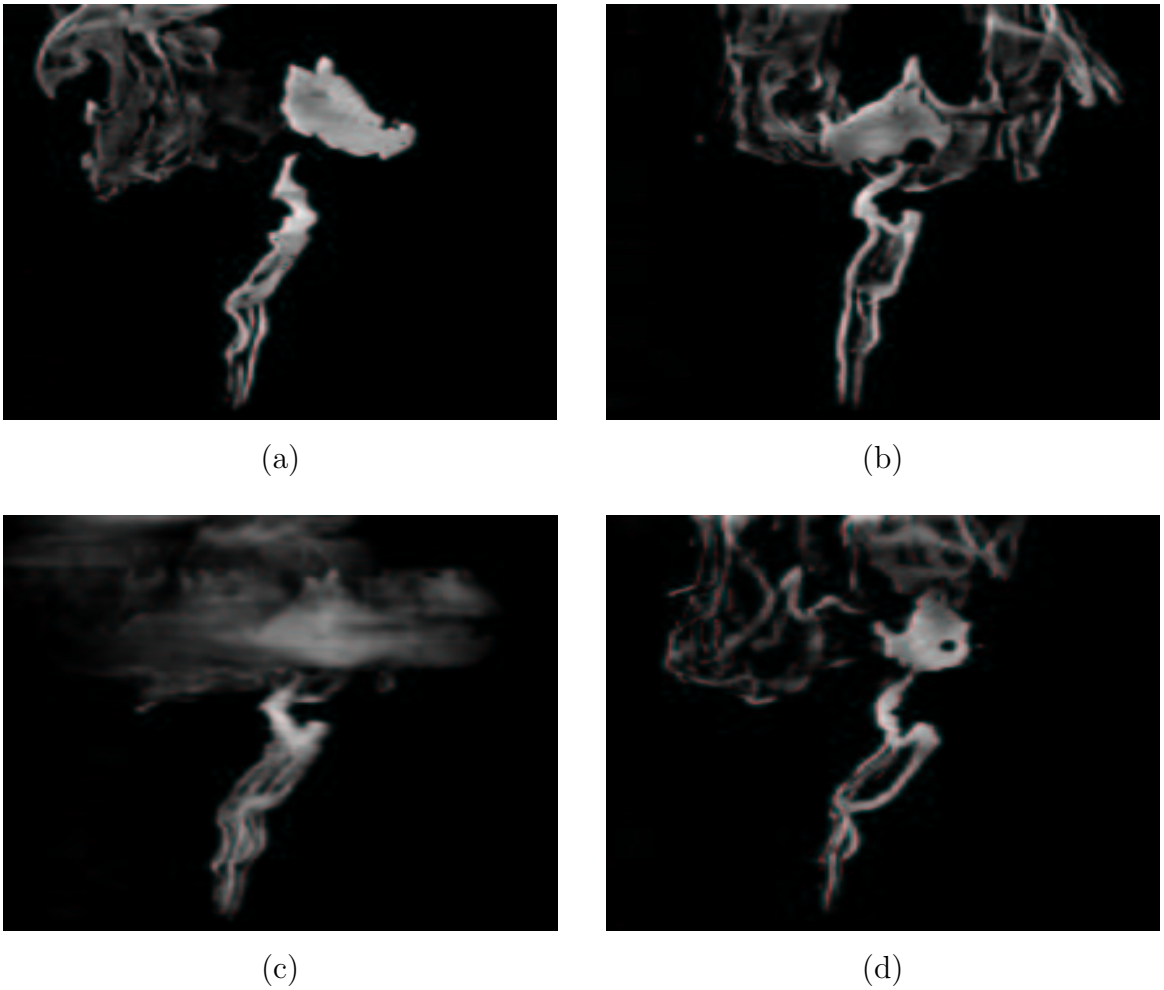
Figure 5.4(b) illustrates that the multiplication solution generates square-like regions in the density field where the image intensities are roughly constant. However, this effect is much more noticeable from top-down or oblique views.

The multiplication solution is shown applied to the MATRIX dataset in Figure 5.5. These results support our intuition that the multiplication solution, as the maximum entropy solution, is the most spread out solution possible still projecting to the input images.

Theorem 5.2.4 (Sheet solution) *If X and Y are input images corresponding to the row and column sums of the density field respectively, there exists a solution to the reconstruction problem S_{sheet} that is zero everywhere except the discrete monotonic curve (x_t, y_t), with $x_t \leq x_{t+1}$ and $y_t \leq y_{t+1}$. This solution is a (locally minimum) low entropy reconstruction.*

Proof We show the existence of S_{sheet} by induction. First, given some $n \times n$ density field M, define the residual images $\hat{X} = X - M\mathbf{1}$ and $\hat{Y} = Y - M^T\mathbf{1}$. Next, define an (i,j)-weak sheet solution as a non-negative density field that

- is zero everywhere except the discrete monotonic curve (x_t, y_t), with $x_t \leq x_{t+1} \leq i$ and $y_t \leq y_{t+1} \leq j$,
- projects to the subimages $X_{1:(i-1)}$ and $Y_{1:(j-1)}$, and
- projects to at least one of $\{X_i, Y_j\}$ with $\hat{X}_i, \hat{Y}_j \geq 0$.



(a) (b)

(c) (d)

Figure 5.5: Multiplication solution compared with ground truth for the MATRIX dataset. (a)-(b) the two input images, (c) 45-degree view of the multiplication solution, (d) ground truth 45-degree view. Note that the multiplication solution by definition reproduces the input images in (a) and (b) exactly.

A base case is the $(1, 1)$-weak sheet solution given by the matrix M that is zero everywhere except $M_{11} = \min(X_1, Y_1)$.

Now we show that any (i, j)-weak sheet solution M can be expanded into a larger weak sheet solution for $i, j < n$. For example, given $\hat{X}_i = \hat{Y}_j = 0$, setting $M_{(i+1)(j+1)} = \min(X_{i+1}, Y_{j+1})$ will yield an $(i + 1, j + 1)$-weak sheet solution. The monotonic shape of the curve is clearly preserved, since row $M_{(i+1)\cdot}$ and column $M_{\cdot(j+1)}$ were previously all zero. Moreover, by the inductive hypothesis and the fact that $\hat{X}_i = \hat{Y}_j = 0$, the expanded solution will project to the subimages $X_{1:i}$ and $Y_{1:j}$. The definition of $M_{(i+1)(j+1)}$ also ensures that M will project to at least one of $\{X_{i+1}, Y_{j+1}\}$ and that all residuals remain non-negative. Thus M has been modified to obtain an $(i + 1, j + 1)$-weak sheet solution.

Now by similar reasoning, if $\hat{X}_i > 0$, we can obtain an $(i, j+1)$-weak sheet solution by setting $M_{i(j+1)} = \min(\hat{X}_i, Y_{j+1})$. Furthermore, if $\hat{Y}_j > 0$, we can obtain an $(i+1, j)$-weak sheet solution by setting $M_{(i+1)j} = \min(X_{i+1}, \hat{Y}_j)$.

Next, consider the (n, j)-weak sheet solution M. If $\hat{X}_n = 0$, then $\sum X = \sum Y_{1:j-1} + (Y_j - \hat{Y}_j) = \sum Y$ because image X is fully explained by the density field and both images have the same sum. This implies that $\hat{Y}_j = 0$, so M must already be a full sheet solution. Alternatively, if $\hat{Y}_j = 0$ (and $j < n$) then we can expand M as before to obtain an $(n, j+1)$-weak sheet solution. By a symmetric argument, we can expand any (i, n)-weak sheet solution or show it is already a full sheet solution. Thus, by a series of expansions from a (1,1)-weak sheet solution we can always obtain a full $n \times n$ sheet solution.

Due to the monotonic shape of the curve, any rectangle aligned to the grid in S_{sheet} will have a density of zero in at least one corner. Then any switching operation applied to such a rectangle can only spread out the mass and increase the total entropy. Thus, the entropy for the sheet solution is a local minimum. ∎

Intuitively, we can visualize starting with one of the 1D input images as the curve, then spreading the densities out (preserving the monotonic shape) so that the curve projects to the other input image as well. This spreading will always be possible because both images sum to one. As illustrated in Figure 5.4, there are in fact *two* such sheet solutions, corresponding to the two directions of monotonicity.

For an implementation of one kind of sheet solution, see Algorithm 1. Note that the sheet solution can also be obtained by iteratively applying switching operations to the multiplication solution, arbitrarily preferring one diagonal orientation to the other, until no further switching operations are possible.

While in some sense, all solutions that project to the input images are equally valid, different solutions can have vastly different effects on image synthesis. Figure 5.6 illustrates the dramatic difference between the multiplication and sheet solutions for the MATRIX dataset. While the multiplication solution doubles impulses, creating halos, the sheet solution might ultimately be perceived as a transparent surface rather than a 3D volume.

As suggested earlier by Figure 5.3, a prominent effect of the multiplication solution is the doubling of sharp features. Consider a density field consisting of two 2D box functions in general position, so that both images consist of two 1D box functions. The reconstruction found using the multiplication solution will then consist of four 2D box

Algorithm 1: SHEETSOLUTION(X,Y)

input : input images X and Y corresponding to row and column sums
output : density field M consisting of the sheet solution S_{sheet}

$i \leftarrow 1$
$j \leftarrow 1$
$M \leftarrow \mathbf{0}_{n \times n}$
while $i \leq n$ *and* $j \leq n$ **do**
  if $X(i) \leq Y(j)$ **then**
    $M(i,j) \leftarrow X(i)$
    $Y(j) \leftarrow Y(j) - X(i)$
    $i \leftarrow i + 1$
  else
    $M(i,j) \leftarrow Y(j)$
    $X(i) \leftarrow X(i) - Y(j)$
    $j \leftarrow j + 1$
return M

functions at the corners of a rectangle.

5.2.3 General Space of Reconstructions

We now expand the analysis to more than two input images and discuss how additional input images can reduce the space of solutions that projects exactly to the input images.

Theorem 5.2.5 *The space of density fields that project to the input images is convex. In other words, given any two density fields S_1 and S_2 that project to the input images, $\lambda S_1 + (1-\lambda)S_2$ will also be a solution for any $0 \leq \lambda \leq 1$.*

Proof Any particular discrete line integral corresponding to some image pixel must be the same for all density fields that project to the input images. Thus any convex combination of density fields that project to the input images will also project to the input images. Note that the restriction on the domain of λ is sufficient to ensure that the new density field will be non-negative everywhere. ∎

Now for $k > 2$ input images, the reconstruction must lie in the intersection of solution spaces defined by all $\binom{k}{2}$ pairs of input images. As described by Theorem 5.2.1, any pair

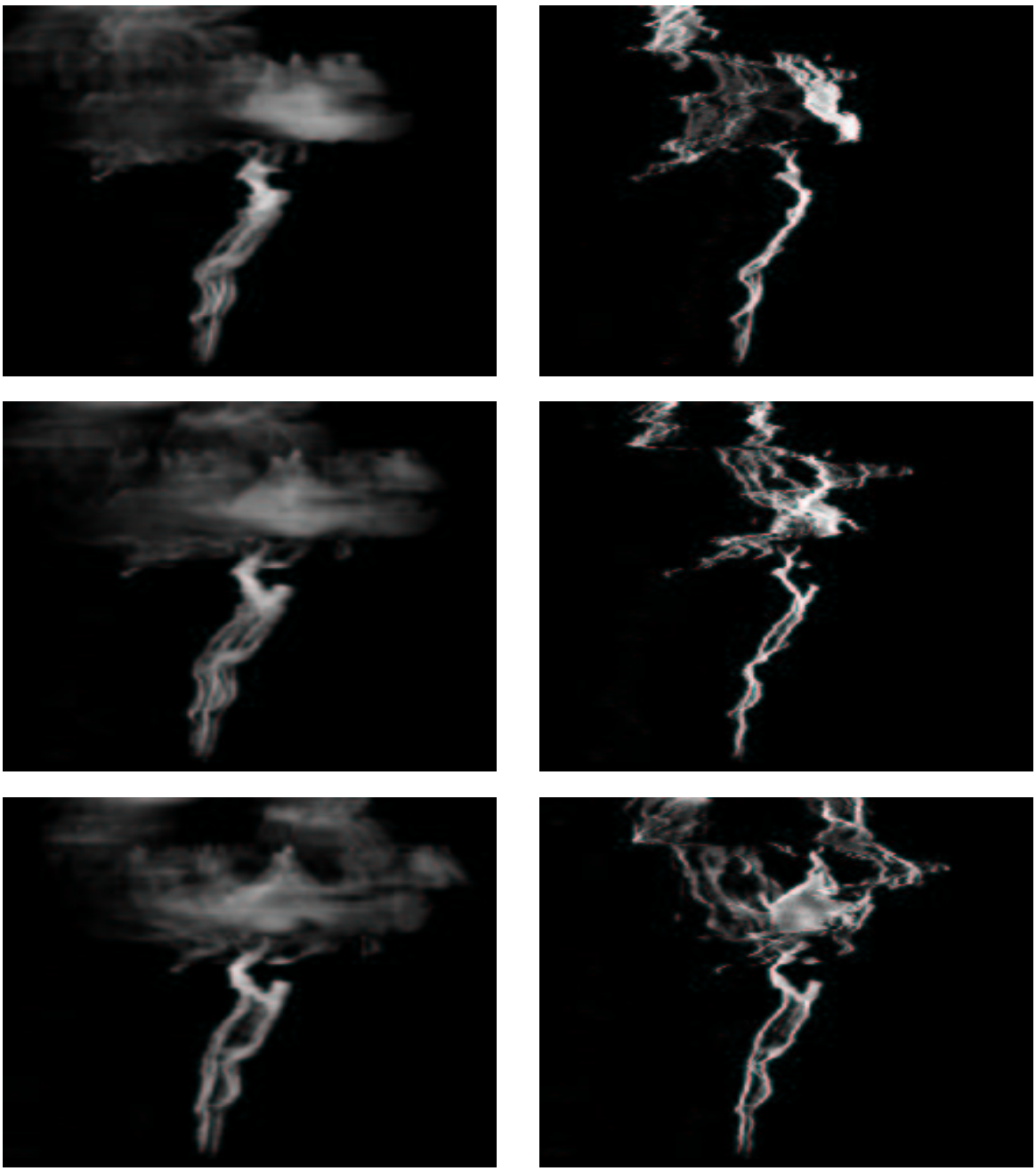


Figure 5.6: Multiplication solution (left) compared with the sheet solution (right) for intermediate views of the MATRIX dataset. The rows correspond to a sequence of three intermediate views. Both solutions reproduce the input images exactly.

of input images restricts the space of solutions to some $(n-1)^2$-dimensional hyperplane lying in the vector space of $n \times n$ matrices.

As more views are added we expect some of the ambiguity to be resolved. In particular, the addition of a third view is expected to eliminate much of the ambiguity associated with the doubling of sharp features found in the multiplication solution (see Figure 5.3). However, additional views are less helpful for disambiguating smooth, diffuse regions. As in standard stereo, the lack of high frequency detail in these regions makes unambiguous reconstruction more difficult.

Ideally, when reconstructing fire, we would like to stay within the space of photo-consistent reconstructions and select the best reconstruction in terms of its plausibility as fire. In theory, optimization methods could be applied to maximize some fire plausibility function over the space of $n \times n$ matrices, constrained to the linear subspace consistent with the input images. However, the high dimensionality of the problem is very restrictive, so for reasons of computational efficiency we pursue a different approach.

5.3 Image-Based Reconstruction Method

We develop an image-based algorithm for the reconstruction of fire that allows us to render the fire from novel points of view. The fire is reconstructed in terms of a density field parameterized by a superposition of blobs. We show that by fixing the positions and variances of the blobs, the reconstruction is reduced to solving a constrained linear least-squares problem. Next, we describe how to improve our initial selection of the positions and variances of the blobs using a stochastic resampling technique based on the previous reconstruction results.

Blobs have been previously used in various fields for both the representation and reconstruction of 3D volumes. For example, parametric models of several Gaussian blobs have been used in the atmospheric sciences to reconstruct gas distributions from very sparse line-integral data [35, 52]. In other work, Marabini *et al.* recently extended the algebraic reconstruction technique (ART) to use 3D blobs as basis functions, and showed that this method is a quantifiable improvement over standard filtered backprojection [69]. As well, Stam simulated and rendered fire using 3D blobs initialized as spherically-symmetric Gaussians, but that become warped over time by diffusion and advection [110]. While spherical Gaussian blobs may be a poor choice in 3D to represent the detailed wispy geometry often found in flames [110], this criticism is less relevant when

the fire in considered in 2D cross-section.

The objective function used is simple least-squares agreement with the input images. Other desirable properties of fire such as coherence are specified implicitly in the smooth nature of the blobs, the limited number of blobs used, and their parameterization.

5.3.1 Parameterized Density Representation

Our approach is to parameterize the density field with a superposition of spherically-symmetric Gaussian blobs. This has two beneficial effects: reducing the dimensionality of the reconstruction problem and implicitly enforcing spatial coherence. Symmetric Gaussian blobs are especially useful because they can be described by relatively few parameters (position, variance, and weight) and they have simple analytic properties. Other 2D parameterizations are also conceivable, including general elliptical Gaussians, other radial basis functions, and wavelets.

A normalized symmetric 2D Gaussian centred at (μ_x, μ_y) is given by the well-known equation:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}\right),$$

where σ^2 is the variance in all directions. The density field can then be represented as a (non-negative) weighted sum of r blobs:

$$\sigma_t \rho(x, y) = \sum_{i=1}^{r} w_i G_i(x, y).$$

For notational simplicity we omit σ_t and incorporate this scalar into the weights $\{w_i\}$.

5.3.2 Linear Formulation

By fixing the positions and variances of the blobs, we restrict the problem to finding the blob weights such that the corresponding density field reproduces the input images as accurately as possible. As we show, this yields a linear system that can be solved in a least-squares sense subject to non-negativity constraints.

First note that line integrals over symmetric Gaussian blobs have an especially simple analytic formula. Defining $\Gamma(G, \ell)$ as the line integral of blob G over the line ℓ, we have:

$$\begin{aligned} \Gamma(G, \ell) &\triangleq \int_\ell G(x, y)\, \mathrm{d}x\mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-d^2}{2\sigma^2}\right), \end{aligned} \tag{5.8}$$

where d is the perpendicular distance from the line to the centre of the Gaussian blob. If the line ℓ is given in implicit form by $\ell_1 x + \ell_2 y + \ell_3 = 0$ then we have:

$$d = \frac{|\ell_1 \mu_x + \ell_2 \mu_y + \ell_3|}{\sqrt{\ell_1^2 + \ell_2^2}}.$$

The geometry of this line integral is illustrated in Figure 5.7.

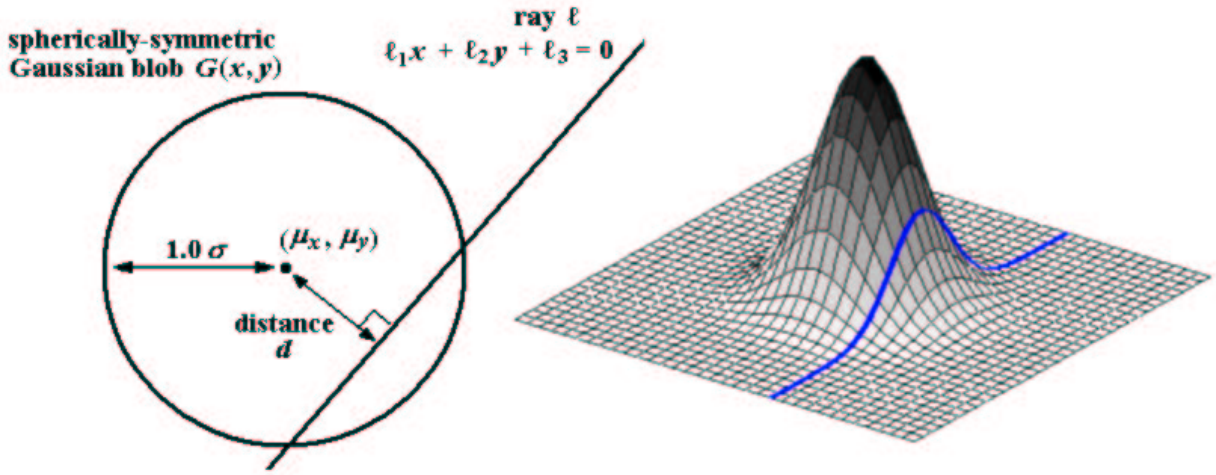


Figure 5.7: Line integral through a symmetric 2D Gaussian blob. Shown as a diagram (left) and illustrated by a 3D plot (right).

From our image formation model of fire in Equation (5.4), the transformed image intensities are related to the corresponding line integrals according to $I'(\ell) = \sigma_t \int_\ell \rho(x, y)\, \mathrm{d}x\mathrm{d}y$. Then from our previous definition in Equation (5.8), for a density field consisting of a single normalized Gaussian blob $\sigma_t \rho(x, y) = G(x, y)$, we have that $I'(\ell) = \Gamma(G, \ell)$. This generalizes readily for a density field given as a linear superposition of blobs:

$$\begin{aligned} I'(\ell) &= \int_\ell \left(\sum_{i=1}^{r} w_i G_i(x, y) \right) \mathrm{d}x\mathrm{d}y \\ &= \sum_{i=1}^{r} w_i \Gamma(G_i, \ell). \end{aligned} \tag{5.9}$$

This represents the complete forward projection model for a density field parameterized by r symmetric 2D Gaussian blobs.

Now given the positions and variances of the r blobs, as well as q pixels in total from the 1D input images, we can build a linear system of equations with blob weights as the unknowns. Each input pixel gives an equation in the form of Equation (5.9), yielding q

equations in the blob weights. In this system, the vector b contains the q transformed pixel intensities, the vector x consists of the r unknown blob weights, and the relatively sparse $q \times r$ matrix A encodes both the camera model and the positions and variances of the blobs.

If $q \geq r$ and $\text{rank}(A) = r$, then the system is overdetermined and standard least-squares techniques can be used to find the solution minimizing the squared image error $\|Ax - b\|_2$. On the other hand, if $q < r$, then the system is underdetermined, and there is an infinite space of solutions with zero residual. Applying the pseudo-inverse least-squares method to an underdetermined system will find the solution in the zero-residual solution space with minimum norm. In practice, to avoid overfitting and for implementation reasons (see §5.3.5) we restrict the number of blobs so that the system is always overdetermined.

We must also be careful to impose the constraint that the density field is non-negative everywhere. Naïvely applying a least-squares fit will result in a solution that is both over-fit and non-physically negative in many areas. Thus, we enforce the sufficient condition that all of the blob weights $\{w_i\}$ are non-negative, yielding the bound-constrained least-squares problem,

$$\min_x \|Ax - b\|_2, \text{ subject to } x \geq 0.$$

This is a special case of the quadratic programming (QP) problem [85]. In fact, constrained least-squares problems of this form are known to be convex, so finding the global minimum is theoretically possible.

5.3.3 Sampling the Space of Blob Configurations

Intuitively, determining the blob positions and variances has a pivotal role in the quality of the reconstruction. For example, if all the blobs are placed in an area that contains zero fire density in the true scene, the errors could be arbitrarily large, even though better placement could yield almost perfect reconstructions under certain conditions. In general, we would like to shift blobs from empty areas to areas containing fire density, and from diffuse areas to areas of fine geometric detail. The challenge is that we do not know where these areas are in advance.

Since every configuration of blob positions and variances leads to some reconstruction, performing a reconstruction can be thought of as sampling the space of blob configurations. So, we would like to improve the reconstruction quality by modifying the configu-

ration of blobs based on previous reconstruction results. Starting from the reconstruction obtained from some reasonable initial blob configuration, we develop two stochastic methods for repositioning the blobs. Note that the sequence resample-then-reconstruct can be repeated as many times as desired, however in practice we only resample once.

Initially, assuming no prior knowledge of the fire density field, we tile the blobs uniformly in a triangular pattern over the region of interest. This pattern (see Figure 5.8) has the nice property that blobs are an equal distance from their immediate neighbours. We set the variances of the blobs σ^2 so that the distance from a blob to its immediate neighbours is $k\sigma$ for some constant k. Because the number of blobs is limited, the reconstruction obtained from this blob configuration is necessarily coarse.

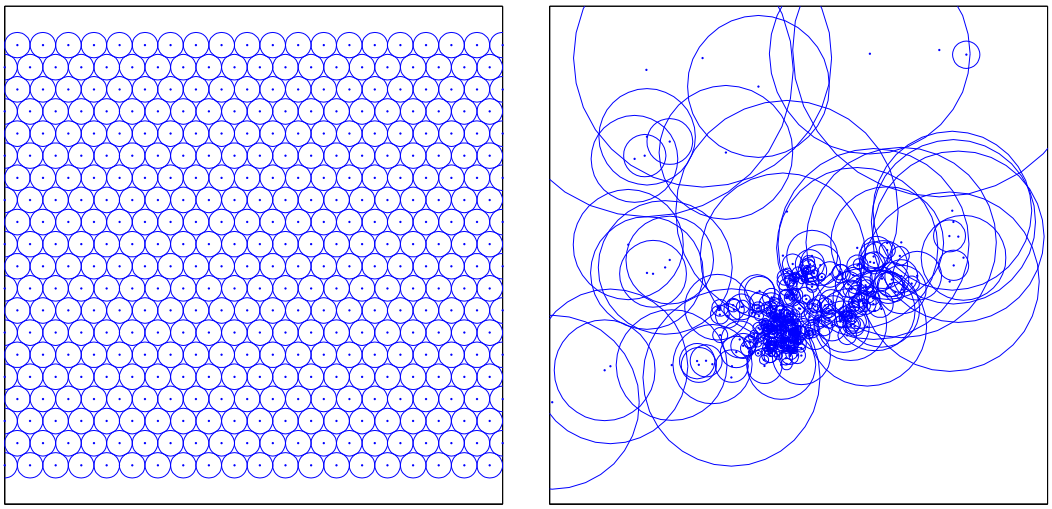


Figure 5.8: Coarse uniform sampling pattern and an example resampled pattern.

The first resampling method we mention is based on the idea that blobs should be placed in regions where there is a high likelihood of non-zero densities. By normalizing the density field from a previous reconstruction, the density field can be treated as a probability function and sampled from directly to obtain new blob positions.

An alternative resampling method is based on the fact that the presence of high frequencies in a region of the density field can only be reproduced accurately if the field is sampled densely enough in that region. Thus, we can resample blob positions according to the local variance of the reconstructed density field.

These blob resampling methods are related to previous work on non-uniform sampling for progressive image compression [40, 91]. This work on image compression also involves

sampling according to a variance-like function, however this is also balanced with another heuristic to ensure that samples are spaced as far apart as possible.

After the positions of the resampled blobs are chosen, their variances can be determined using the same rule as in the initial uniform sampling. An additional consideration, however, is that the variances should be clamped to some reasonable minimum so that blobs project to at least half the width of one image pixel. Otherwise a blob may be so narrow that only one ray from one viewing direction ever intersects it. As shown in Figure 5.10(b), this overfitting manifests itself as speckle noise when images are reconstructed from novel viewing angles.

An improved method of choosing variances involves constructing the Delaunay triangulation (see Figure 5.9) and choosing the variances in accordance with the distance to the furthest neighbour. In this way, the entire density field will be "covered" by the blobs. Specifically, every point in the density field is then guaranteed to lie within k standard deviations of some blob. When this is not the case, as is possible with the nearest-neighbour approach, some region of actual fire density may not be covered by any blob and this density will necessarily be shifted to the wrong location in order to compensate.

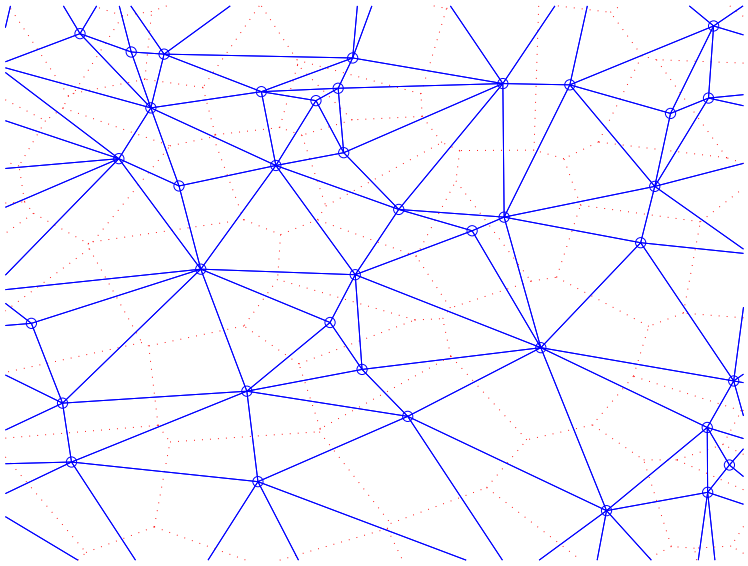


Figure 5.9: Delaunay triangulation (solid) shown over the Voronoi diagram (dotted) for a sample collection of points.

5.3.4 Incorporating Spatial Coherence

Although we split the reconstruction volume into parallel slices for computational efficiency, we still expect nearby slices to exhibit a strong degree of coherence. Specifically, the image information from nearby slices provides joint constraints that reduce the space of photo-consistent reconstructions. We describe two complementary methods for enforcing coherence between nearby slices. In practice, the reconstructions incorporating these methods better capture the structure of the true fire density.

Our first insight about coherence is that the line integrals should project closely to the corresponding image pixels in adjacent slices as well. Thus we define an image smoothing operation over the pixel intensities from the $2d+1$ neighbouring slices. So for slice z, we can solve the smoothed system,

$$A_z x_z = \sum_{i=-d}^{+d} k_i b_{z+i}, \text{ subject to } x_z \geq 0, \tag{5.10}$$

where A_z is the matrix encoding the camera model and the blob configuration, x_z consists of the unknown blob weights, and b_{z+i} contains the pixel intensities for slice $z+i$. The kernel $\{k_i\}$, where $\sum k_i = 1$, characterizes the type of and degree of smoothing.

The smoothed system is equivalent to the following stacked linear system over the $2d+1$ adjacent slices:

$$\begin{pmatrix} \sqrt{k_{-d}}\, A_z \\ \vdots \\ \sqrt{k_{+d}}\, A_z \end{pmatrix} x_z = \begin{pmatrix} \sqrt{k_{-d}}\, b_{z-d} \\ \vdots \\ \sqrt{k_{+d}}\, b_{z+d} \end{pmatrix}. \tag{5.11}$$

Note that we can also construct the full coupled version of this system in which blobs are given true vertical extent, and so contribute intensity to nearby slices. The problem with doing this is that the resulting system is too large to solve efficiently with constrained least-squares methods. Indeed, this was the reason we considered the slices separately in the first place.

Another opportunity to exploit coherence is in the resampling step. When performing stochastic resampling, we can actually resample from adjacent slices as well. For example, we can resample with equal probability from the current slice, the slice above, and the slice below.

5.3.5 Algorithm and Implementation Details

In this section, we give the overall algorithm (see Algorithm 2) and describe some of the specific implementation details. The reconstruction algorithm was implemented entirely in Matlab 6.1 and has not been optimized for speed.

Algorithm 2: FIRERECONSTRUCTION

input : q input pixels $I(\ell_1), \ldots, I(\ell_q)$ corresponding to the lines $\ell_1, \ldots, \ell_q$, blob spacing constant k, number of resampling steps

output : density field parameterized in terms of r blobs, $\rho = \sum_{i=1}^{r} w_i G_i$

1. Transform the intensities according to Equation (5.4) or Equation (5.6) to obtain $I'(\ell_1), \ldots, I'(\ell_q)$. Optionally apply image smoothing with neighbouring slices as described in §5.3.4.

2. Tile the density field uniformly with r Gaussian blobs $G_1, \ldots, G_r$. Set the blob variances σ^2 so that neighbouring blobs are spaced $k\sigma$ apart.

3. Given the positions and variances of the blobs, build a linear system in the r unknown blob weights, with one equation for each of the q input pixels:

$$\begin{pmatrix} I'(\ell_1) \\ \vdots \\ I'(\ell_q) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{r} w_i \Gamma(G_i, \ell_1) \\ \vdots \\ \sum_{i=1}^{r} w_i \Gamma(G_i, \ell_q) \end{pmatrix}.$$

4. Solve this system in the least-squares sense, subject to $w_1, \ldots, w_r \geq 0$. This gives us a new estimate for the density field, $\hat{\rho} = \sum_{i=1}^{r} w_i G_i$.

5. Terminate and return $\hat{\rho}$ if the desired number of resampling steps have been done.

6. Resample the positions of the blobs according to the local variance of the reconstructed density field $\hat{\rho}$. Optionally perform the resampling over neighbouring slices as described in §5.3.4. Set the blob variances σ^2 so the distance to the furthest neighbour according to the Delaunay triangulation is $k\sigma$. Goto step 3.

For both the initial uniform blob tiling and the resampling step, we choose blob variances according to the heuristic that the distance to some neighbour is $k\sigma$. We choose k to be small enough so that the mass is not overly concentrated at the sample

point with empty space surrounding it, but large enough so that the samples can resolve relatively small areas. In practice, we found that $k = 2.0$ was an appropriate value for a wide range of datasets.

We use the `lsqlin` function in the Matlab optimization toolbox to solve the constrained least-squares problem. This function operates using an interior-reflective trust-region method, where conjugate gradient is used to solve a large linear system approximately at each iteration [85]. In practice, `lsqlin` does not converge on the global minimum due to numerical error.

By limiting the number of blobs to no more than the number of image pixels q, we avoid the underdetermined case. This helps reduce overfitting, exemplified by the extreme case where the density field is parameterized so finely that each image pixel projects to its own tiny blob. Another practical issue is that when solving constrained linear systems, Matlab performs several orders of magnitude slower for the undetermined case. The reason for this is that the initial feasible point is found inefficiently by linear programming and the underdetermined algorithm does not take advantage of the sparsity of A. On the other hand, limiting the number of blobs in this fashion necessarily limits the resolution of the reconstruction. As we discussed in §5.3.3, resampling can be used to better configure a limited number of blobs in order to preserve fine geometric detail.

In practice we resample only once, according to the local variance of the estimated density field reconstructed from the coarse uniform blob tiling. We then use the Delaunay triangulation to set the variances according to the furthest neighbours.

We also implement the coherence techniques described in §5.3.4, enforcing coherence with the two directly adjacent slices. For the image smoothing step we use the weights $[0.5, 1, 0.5]$, somewhat deemphasizing the adjacent slices. We also modify the resampling step to resample with equal probability from the current slice and the two adjacent slices.

5.3.6 Experimental Results

For testing purposes we obtained datasets consisting of synchronized images of real and simulated fire. Where possible, ground truth data was obtained to help evaluate the quality of the reconstruction. For real fire, some input images can be withheld from the reconstruction algorithm and later used for validation. For simulated fire, a complete ground truth 3D representation of the fire is given.

MATRIX This dataset was extracted from a promotional video for Digital Air Incor-

porated [32] featuring various demonstrations of their commercial 3D freeze-frame system known as Timetrack. Using Timetrack involves positioning a flexible gantry holding up to a hundred and fifty synchronized cameras around the scene. As seen in films such as *The Matrix* (1999), animating the sequence of still images obtained from these cameras gives the appearance of fluid camera motion through a scene where time has been stopped.

In one segment of the video, Timetrack was applied to a geometrically complex gas flame emerging from a jet. We extracted 47 frames from this sequence in which the camera appears to pan approximately ninety degrees around the flame. For lack of explicit information about camera calibration, camera positions were assumed to be evenly spaced on a ninety degree circular arc. The camera model assumed is a hybrid between 2D perspective projection in the planes defined by the scanlines and parallel projection in the vertical direction. The standoff distance of the camera and the focal length were also estimated by hand.

Because the conversion from NTSC video introduced interlacing artefacts, only every second scanline of the MATRIX sequence was used. The horizontal resolution was also downsampled for speed, giving 47 input images with a resolution of 159×213. Most experiments used greyscale (luminance) versions of the input images, ignoring the available RGB information.

PLUME This dataset consists of a thick plume of fire simulated using the NIST Fire Dynamics Simulator (FDS 2.0) [84]. The 3D density and temperature fields were obtained at a resolution of $65 \times 65 \times 145$ voxels, several seconds into the simulation. Synthetic input images (65×145) were then produced based on the density field using a parallel projection model.

HOOD This is another dataset obtained from simulation results of the NIST Fire Dynamics Simulator. Here a turbulent jet of flame rises up to meet an unseen ventilation hood. The resolution of this dataset is $97 \times 97 \times 65$ voxels and the synthetic input images (97×65) were produced similarly.

CAMPFIRE This dataset was obtained using the physically-based simulation model for fire recently presented in [82]. It consists of a $90 \times 90 \times 240$ voxel grid storing the density and temperature field of a campfire. Synthetic input images (90×240) were produced based on the density field using a parallel projection model.

TORCH This dataset was obtained in the laboratory by capturing synchronized images of a burning patio torch. Twin Sony DXC-9000 3CCD cameras were used in progressive scan mode to feed synchronized (genlocked) images to two Matrox Meteor II video capture boards. The image resolution was 640×480. Checkerboard patterns visible from both cameras were used for calibration and to calculate the extrinisic parameters of the cameras as well [16]. The captured images were rectified according this calibration so that corresponding scanlines form 1D images of the same epipolar plane.

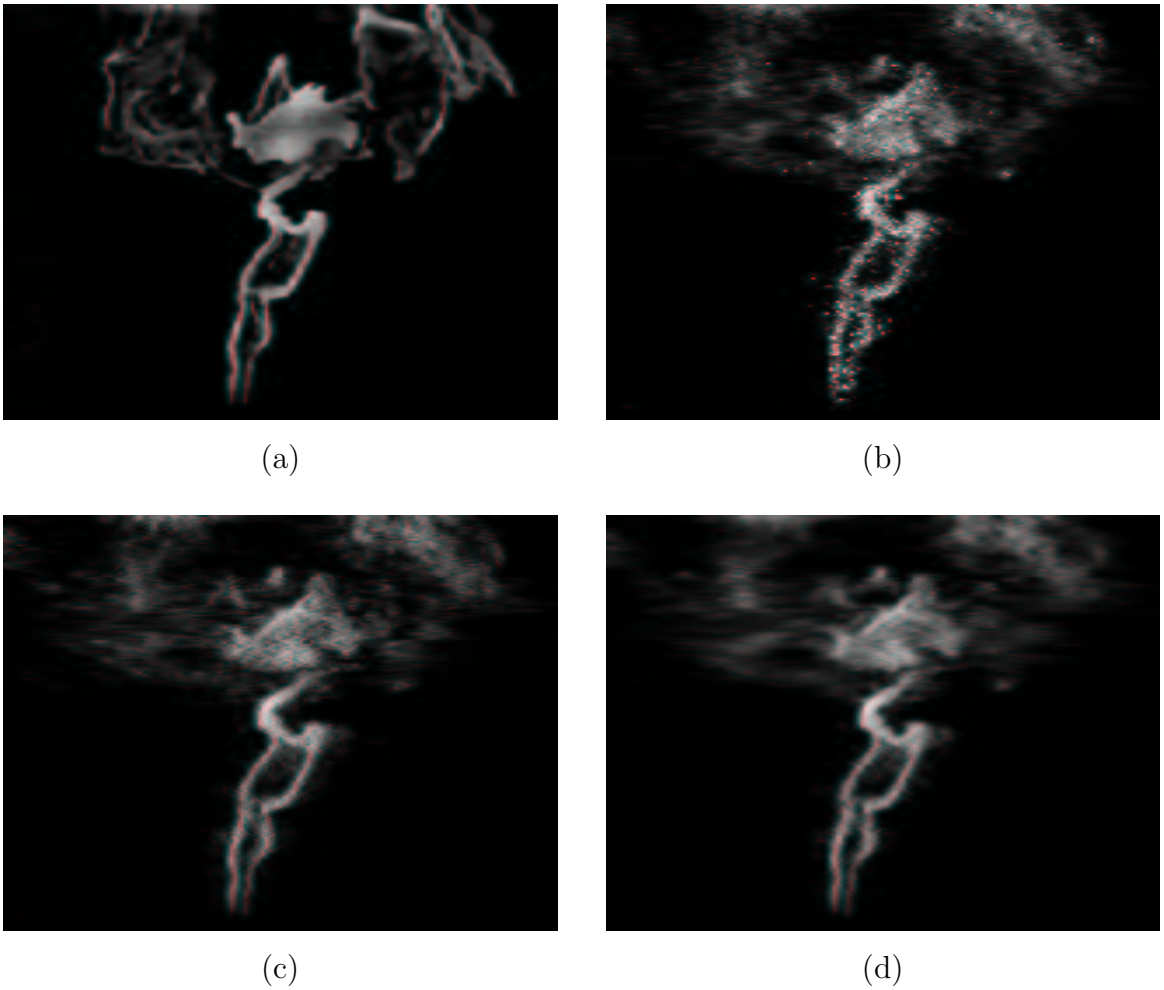
The reconstruction method is relatively slow, with the bottleneck being the constrained least-squares solver. The solver does, however, take advantage of the sparsity of the matrices (typically 5–20% non-zero entries) which gives a several-fold speedup over dense matrices. To give a concrete example, for the MATRIX dataset, solving one slice consisting of 800 input pixels and 400 blobs takes approximately 15 seconds on a PIII-500. Thus a complete fit for the MATRIX dataset takes a total of about 50 minutes.

The different methods of choosing variances are compared for the MATRIX dataset using input images from five evenly-spaced cameras. Numerical results are summarized in Table 5.1, giving per-pixel mean squared error over both the five input images and the entire set of 47 available ground truth images. Reconstruction quality is best assessed by examining error with respect to the ground truth images, which leads us to favour the Delaunay triangulation method. Because the improvements of the Delaunay method were achieved by reducing overfitting, this method unsurprisingly gives the largest error for the input images. As we see in Figure 5.10, using the Delaunay triangulation method gives the best visual result as well.

method of choosing variance	MSE (input)	MSE (all)
nearest neighbour	211.3	539.1
nearest neighbour with minimum	236.8	357.6
Delaunay triangulation	273.1	332.1

Table 5.1: Comparing the mean squared error of reconstructions using different methods of resampling the blob variances for the MATRIX dataset with 5 input images.

As we have already observed, the sequence of resample-then-reconstruct can be performed as many times as desired. However, results show that it is better not to resample more than once, for the sake of both speed and caution. As shown in Figure 5.11 for the MATRIX dataset with five input images, after one resample, the fit to both the input



(a) (b)

(c) (d)

Figure 5.10: Different methods of resampling the blob variances. Reconstructions are shown for an intermediate view of the MATRIX dataset with 5 input images. (a) ground truth, (b) nearest neighbour, (c) nearest neighbour with minimum variance, (d) Delaunay triangulation method.

images and the overall ground truth data actually worsens. This presumably because the local variance in the overfitted resampled reconstructions fails to provide good additional information about where to reposition the blobs.

The coherence techniques described in §5.3.4 were found to give small but consistent improvements, reducing per-pixel mean squared error on the order of 2–5%. Table 5.2 shows the effect of the different coherence techniques for the MATRIX dataset using five input images. As shown, the biggest numerical gain comes from applying both coherence techniques, resampling several slices and smoothing the input images with images from the neighbouring slices. While the coherence techniques yield somewhat more smoothness in the vertical direction, they do not appear to help resolve overfitting.

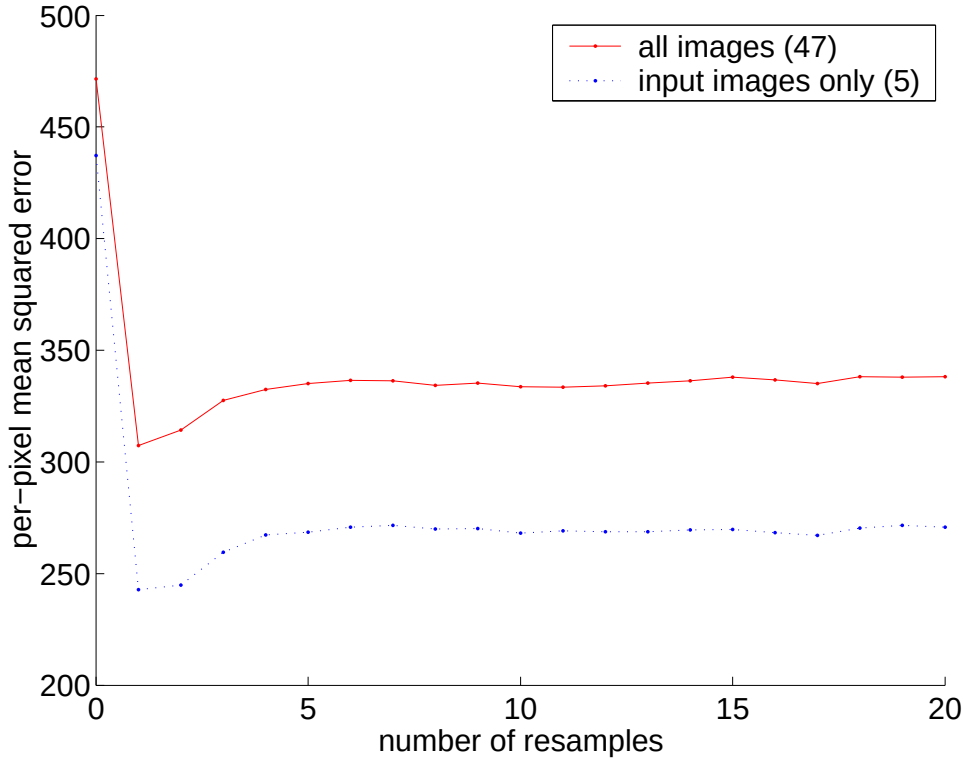


Figure 5.11: Effect of the number of resamples for the MATRIX dataset with 5 input images.

coherence method	MSE (input)	MSE (all)
none	282.6	341.6
sampling	283.4	337.0
image smoothing	270.5	335.6
both	271.5	330.7

Table 5.2: Comparing the mean squared error of reconstructions using different coherence methods for the MATRIX dataset with 5 input images.

We also examined the effect of using different numbers of input images for reconstructing the MATRIX dataset. The error results are plotted in Figure 5.12. As shown in Figure 5.13, the quality of the reconstruction degrades significantly with 2–3 cameras. With so few input images, novel views far from the input views demonstrate serious overfitting artefacts (which are even more striking when the camera is animated). With 4–5 cameras, the results improve considerably, but certain areas of the reconstruction still exhibit ambiguity artefacts and diffuse areas are poorly reconstructed in general. Using 6 or more cameras further reduces some of the ambiguity, but for this dataset

there appears to be a limit on improvement attainable by adding more views, possibly due to the poorly estimated camera calibration.

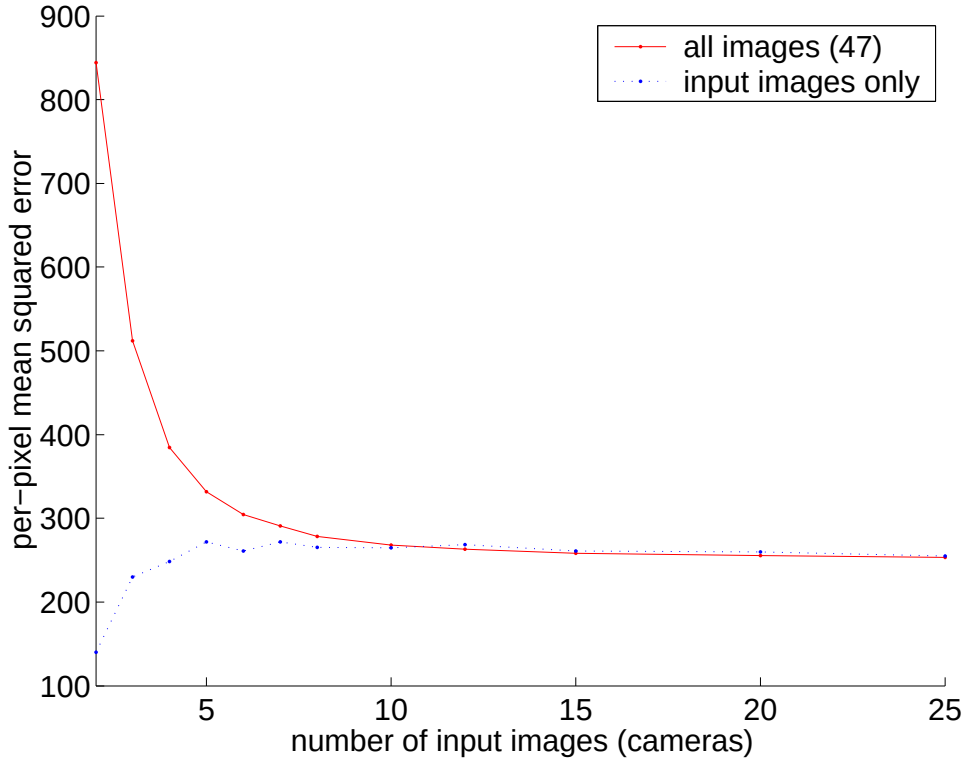
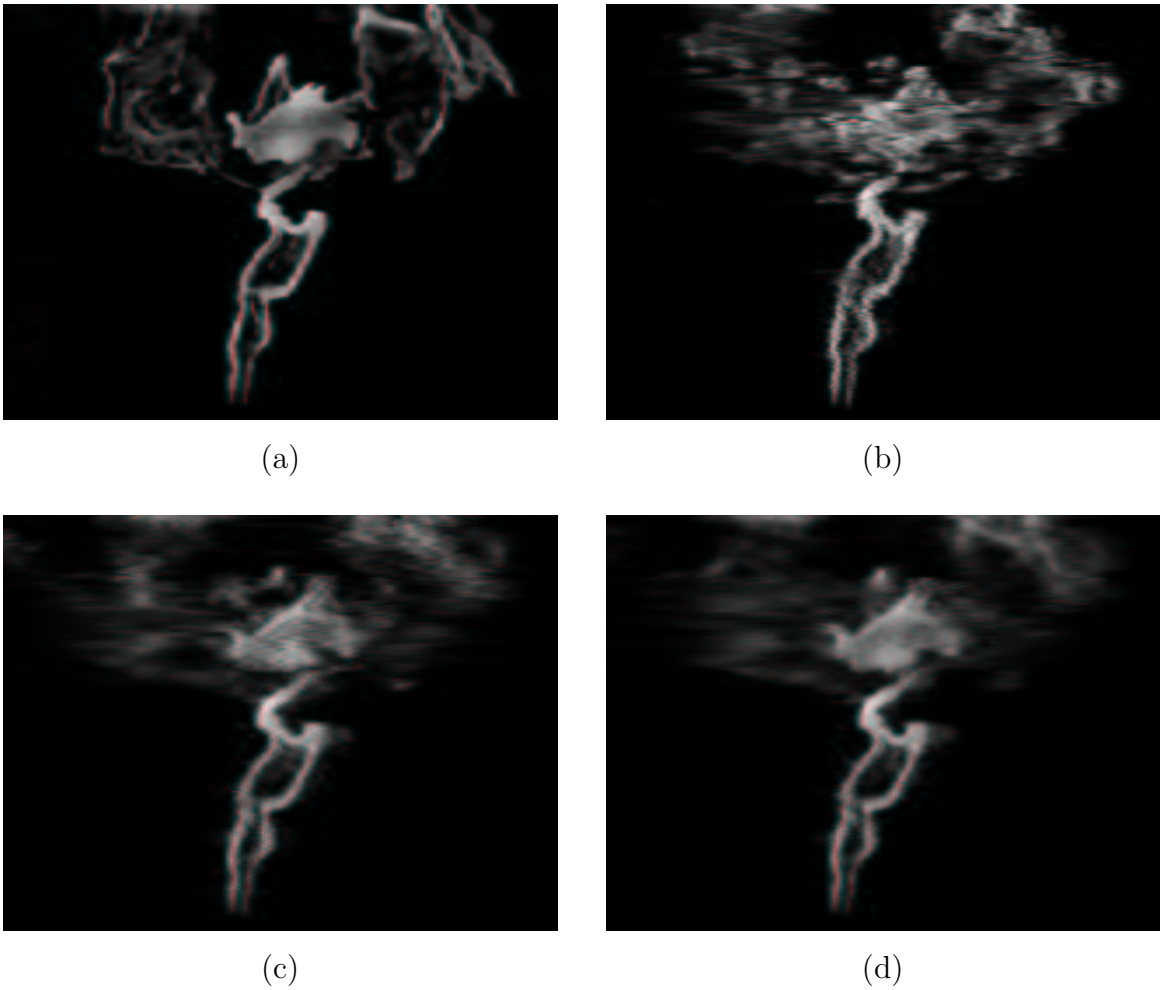


Figure 5.12: Effect of the number of input images for the MATRIX dataset.

Reconstruction results using five cameras are illustrated for the following datasets: MATRIX (Figure 5.14), PLUME (Figure 5.16), HOOD (Figure 5.18), and CAMPFIRE (Figure 5.19). In all cases, the standard reconstruction method described in §5.3.5 was used, where the tunable parameters were left the same for all datasets. Difference images are shown for the MATRIX and PLUME datasets in Figures 5.15 and 5.17 respectively. These illustrate the fact that the errors are not uniformly distributed over the image and highlight the regions of the fire where the reconstruction is poor.

The reconstruction of the the TORCH dataset from two cameras is shown in Figure 5.20. As this dataset was obtained from real fire, no ground truth is available for intermediate views. Nevertheless, away from the input views bad overfitting artefacts are clearly seen, showing the intrinsic ambiguity of the problem and the limitations of the blob-based reconstruction method with only two input views.



(a) (b)

(c) (d)

Figure 5.13: Intermediate view of the MATRIX dataset using different numbers of input images. (a) ground truth, (b)-(d) reconstruction using 2, 5, and 10 cameras respectively.

Ground truth 3D data is compared to slices of the reconstruction for the PLUME dataset (Figure 5.21) and the CAMPFIRE dataset (Figure 5.22). Note that the while the reconstruction preserves some of the major features, the reconstructed slices do not match the structure of the ground truth fire density in general.

In preliminary recent work, it was found that the sheet solution (see §5.2.2) gives good results with only two input views for certain types of fire. To render a sheet solution, we can warp the two input images according to the geometry of the sheet (using 1D image warping) and combine the warped images in a view-based manner. Moreover, colour is easy to incorporate into the sheet solution by calculating the geometry of the sheet according to greyscale versions of the images and warping each of the RGB channels separately. Note that the appropriate sheet solution (of the two) must be selected depending on the viewing direction to avoid rendering the sheet "edge on" as

previously seen in Figure 5.6 (right).

An example of the sheet solution applied to the TORCH dataset is shown in Figure 5.23. Because the transparency of the sheet makes perceiving the underlying 3D shape difficult and the real flame is actually close to sheet-like in shape, in this particular case we obtain a very convincing 3D reconstruction. Another advantage of the sheet solution is its speed to compute. Indeed, we believe that an optimized implementation could compute the sheet solution and render it in real time.

5.3.7 Discussion

With hundreds of input images, it would be possible to accurately estimate the density of the fire over a voxelized space with a resolution comparable to that of the images. However, since the number of input images is severely limited by the number of synchronized cameras available, most techniques that find solutions directly in the voxel domain are not practical. For example, the standard filtered backprojection method (see Figure 4.2) is insufficient for a sparse number of views. Even the algebraic and statistical methods described in §4.1.2–4.1.3 cannot cope with so few (less than six) views. We implemented several of these techniques for testing purposes, including SART [5] and the local tomography method described by Bouman [46], but all of them produced overfitted poor results on such sparse data.

The biggest problem with the proposed reconstruction method is the overfitting of blobs to the input images, especially with only 2–3 input images. In practice it is difficult to appease the conflicting demands that the reconstruction be smooth and fire-like from novel views, yet sharp enough to preserve the fine-scale geometric detail (and spectral characteristics) of the original input images. Because the space of ambiguous reconstructions that project approximately to the input images is very large (see §5.2) the overfitting problem is challenging. We obtain ambiguity artefacts of the sort where some blobby region projects to an obviously "wrong" 3D location, even though all the input images may be faithfully reconstructed in that region. Ideally, we would like to be able navigate the space of solutions for which the input views are reproduced without error, quantifying the plausibility of these solutions in terms of how well they satisfy physically-motivated properties of fire. While coherence may indeed be an intrinsic property of fire, in isolation it is an insufficient measure of plausibility, at least in our implementation.

The physical model of fire is not currently integrated into our objective function

because we obtain only a visual reconstruction of fire. It is likely that if we were able to reconstruct the physical properties of fire, these would allow us to better assess the plausibility of different reconstructions. We might also exploit the temporal aspect of the physical model by considering synchronized video sequences of fire as input. Beyond enforcing the obvious temporal coherence, one might assess a particular reconstructed fire sequence as more plausible if the dynamics were more fire-like.

Intuitively, one might expect that at low enough resolution, the reconstruction should be coarse but not overfit, and thus be more accurate. This suggests the idea of using the solution from coarser levels as constraints to guide the construction of finer level solutions. In practice, however, this idea has not proven very useful. Unfortunately, the overfitting artefacts that emerge even at moderate resolutions tend to lock the algorithm onto poor solutions.

Given a particular configuration of blobs, it is unclear how shifting a blob or changing its variance will affect the reconstruction. Likewise, obtaining a multiscale decomposition of the density field appears difficult because it is unclear where to refine next without performing the reconstruction for the various alternatives and comparing the results. Of course, it is always possible to modify the blob configuration, solve a new constrained linear system, and check whether the new reconstruction has improved. In this sense, we can view the reconstruction method as sampling a point in the high-dimensional space of all blob configurations. However, we would prefer an analytic method for navigating the space of blob configurations that would spare us the high computational cost of solving constrained linear systems as a blind sampling step. Perhaps searching this high-dimensional space of blob configurations would be amenable to more sophisticated randomized sampling methods like Markov-chain Monte Carlo (MCMC) methods [81] or simulated annealing [117].

We might also obtain better results by using more natural shapes for the blobs. For example, the blobs could be initially defined as spherically-symmetric Gaussians for the coarse fit, but then modified during resampling to better fit the reconstructed local variance field. Perhaps we could even apply machine learning techniques to obtain good basis functions for fire from simulation results.

Note that solving for the optimal positions, variances, and weights of the blobs in conjunction represents a much harder non-linear problem than the bound-constrained linear problem currently obtained by fixing the positions and variances of the blobs. Such a problem could be approached using standard non-linear optimization techniques like

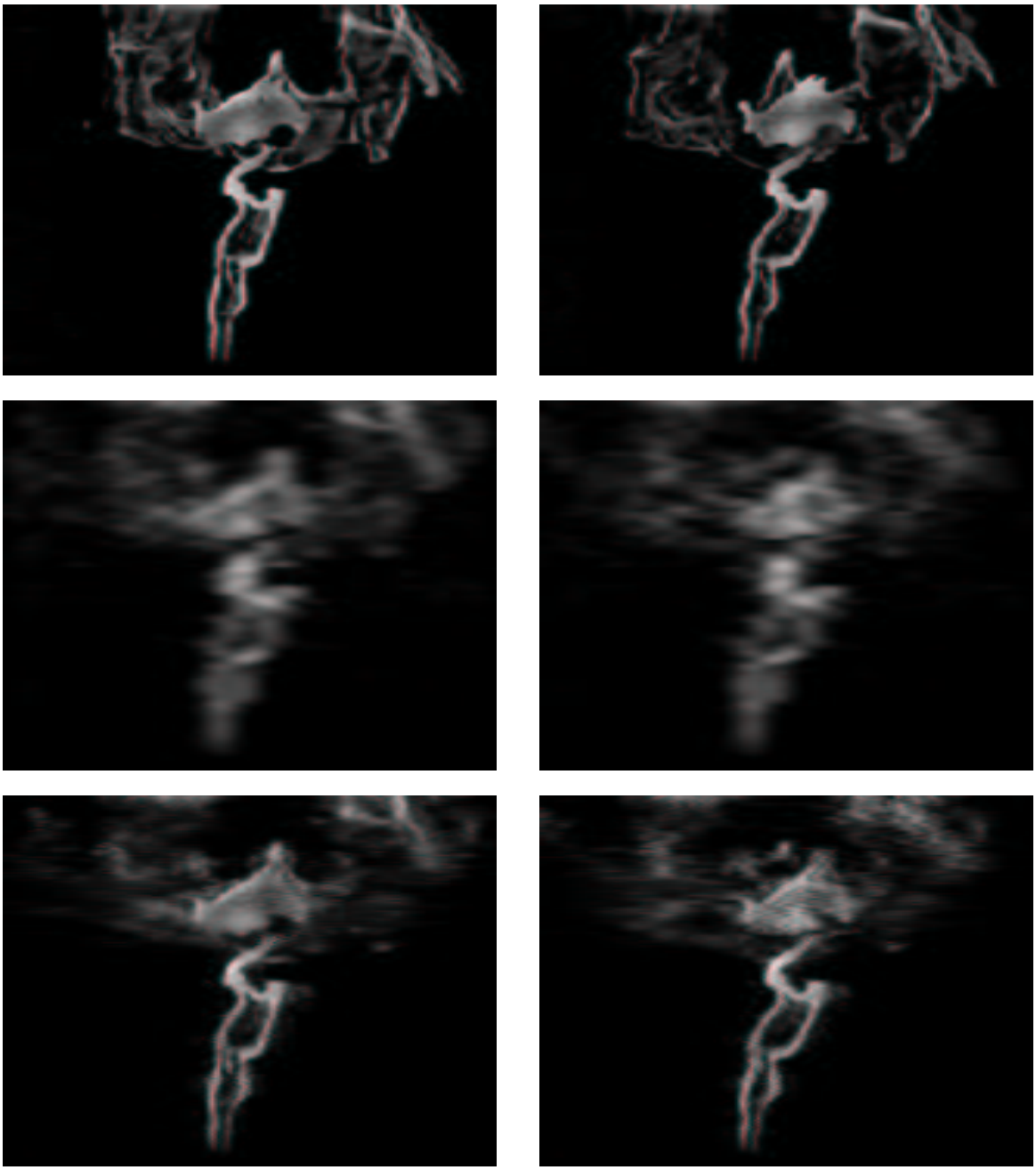


Figure 5.14: Reconstruction of the MATRIX dataset using 5 input images. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. The left column corresponds to an input view and the right column to a novel intermediate view.

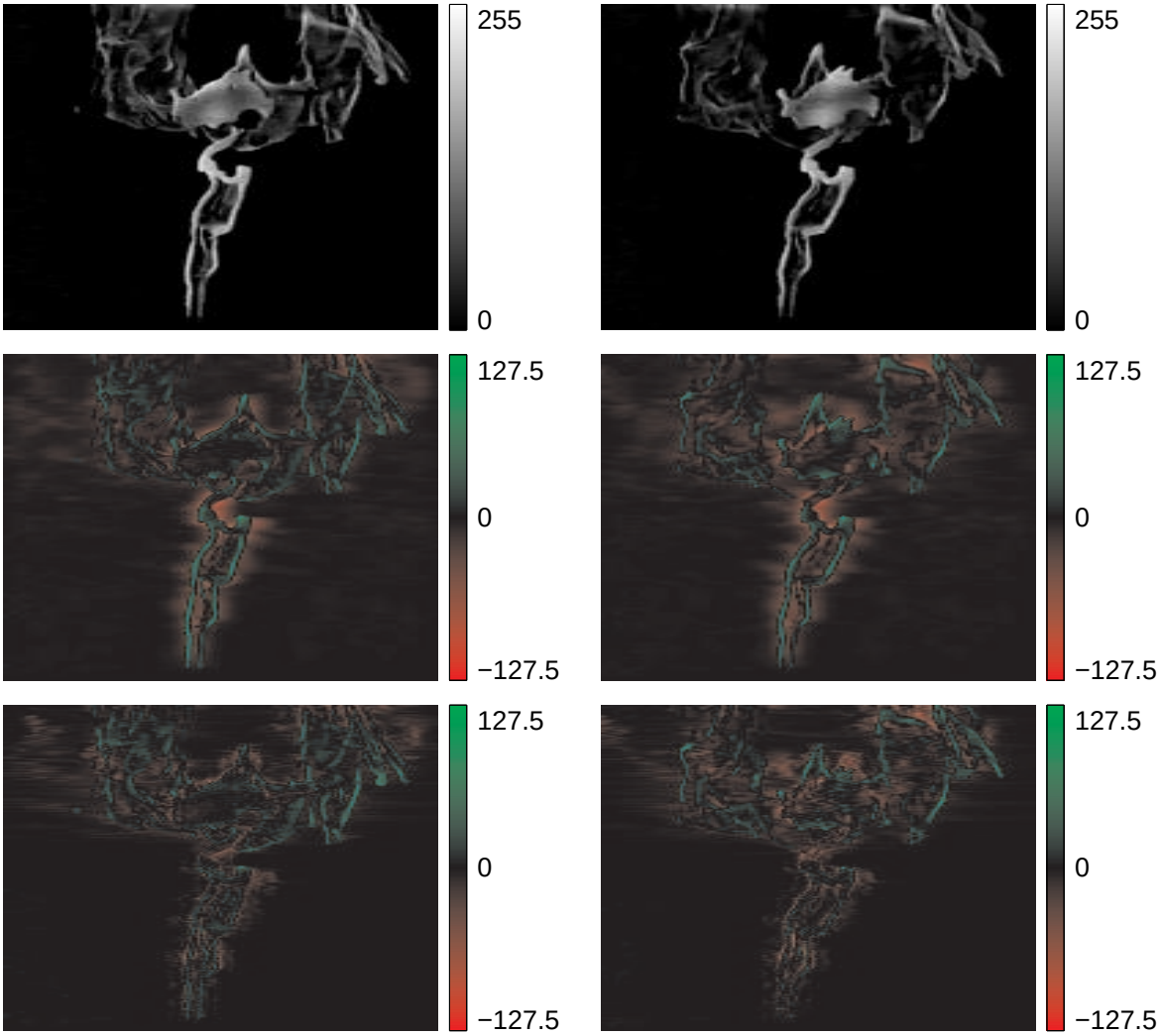


Figure 5.15: Difference images for the MATRIX dataset using 5 input images. From top to bottom: ground truth, difference images for the coarse reconstruction, and difference images for the final reconstruction. The left column corresponds to an input view and the right column to a novel intermediate view. The difference images are shown in false colour.

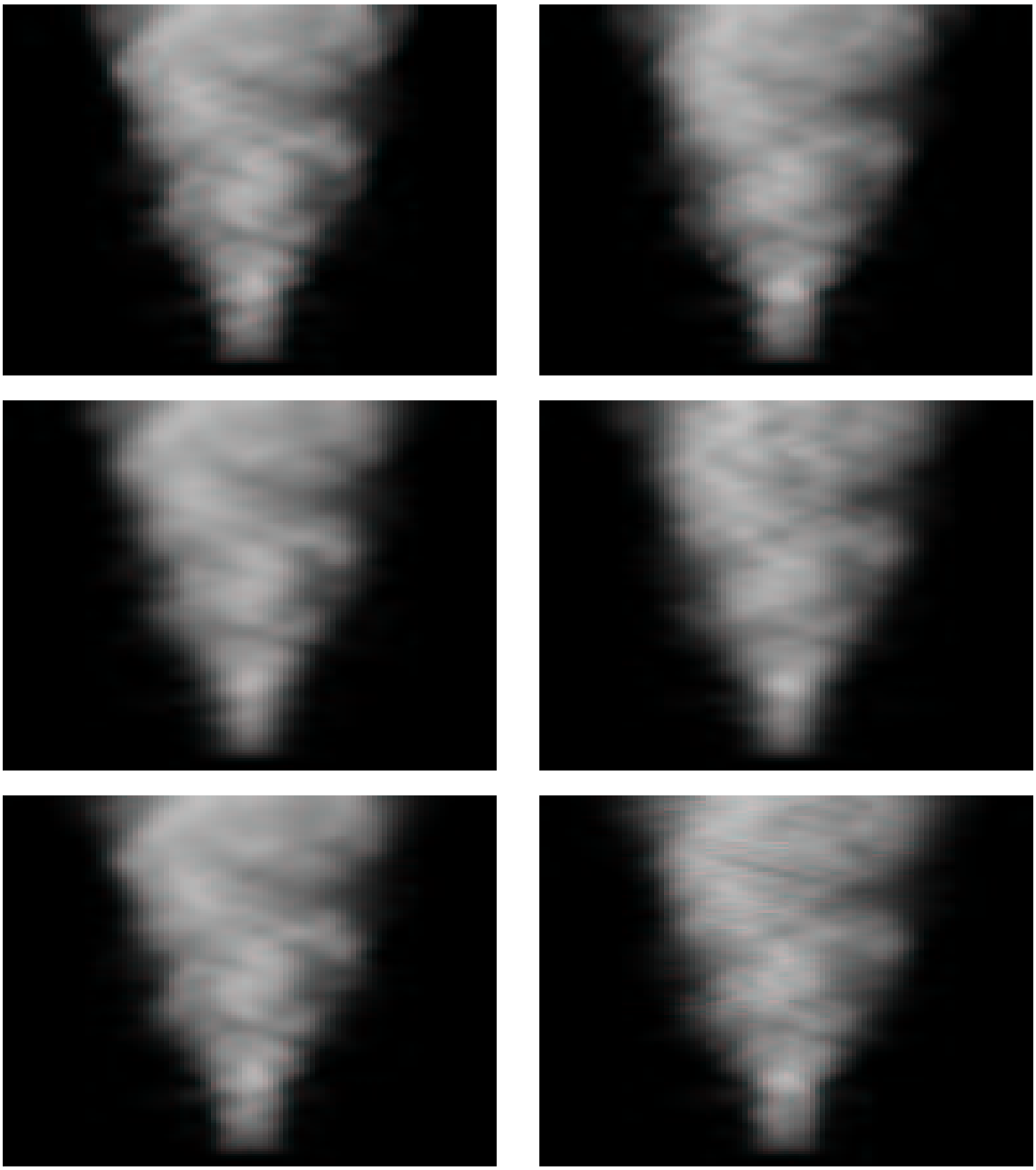


Figure 5.16: Reconstruction of the PLUME dataset using 5 input images. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. The left column corresponds to an input view and the right column to a novel intermediate view.

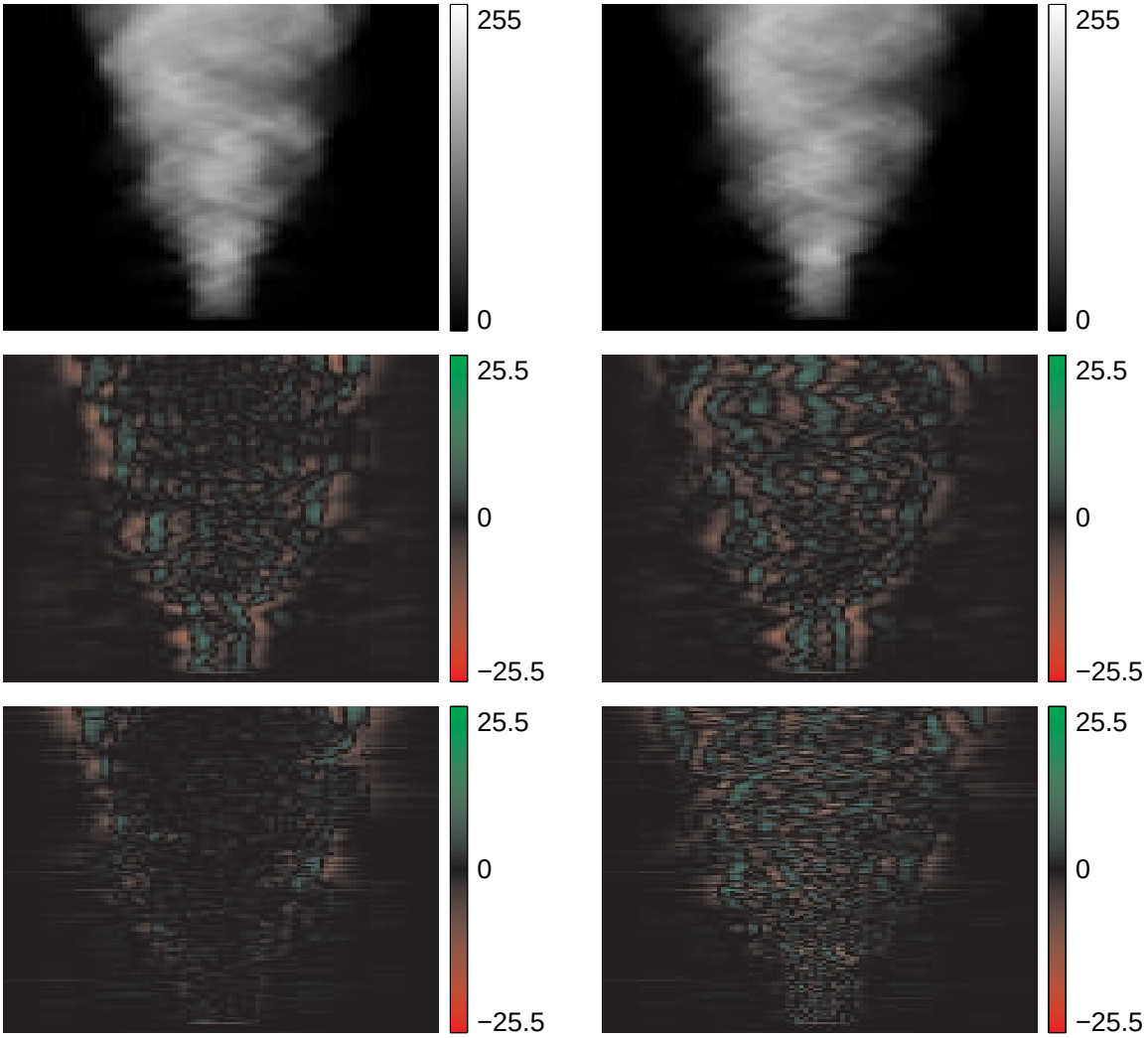


Figure 5.17: Difference images for the PLUME dataset using 5 input images. From top to bottom: ground truth, difference images for the coarse reconstruction, and difference images for the final reconstruction. The left column corresponds to an input view and the right column to a novel intermediate view. The difference images are shown in false colour.

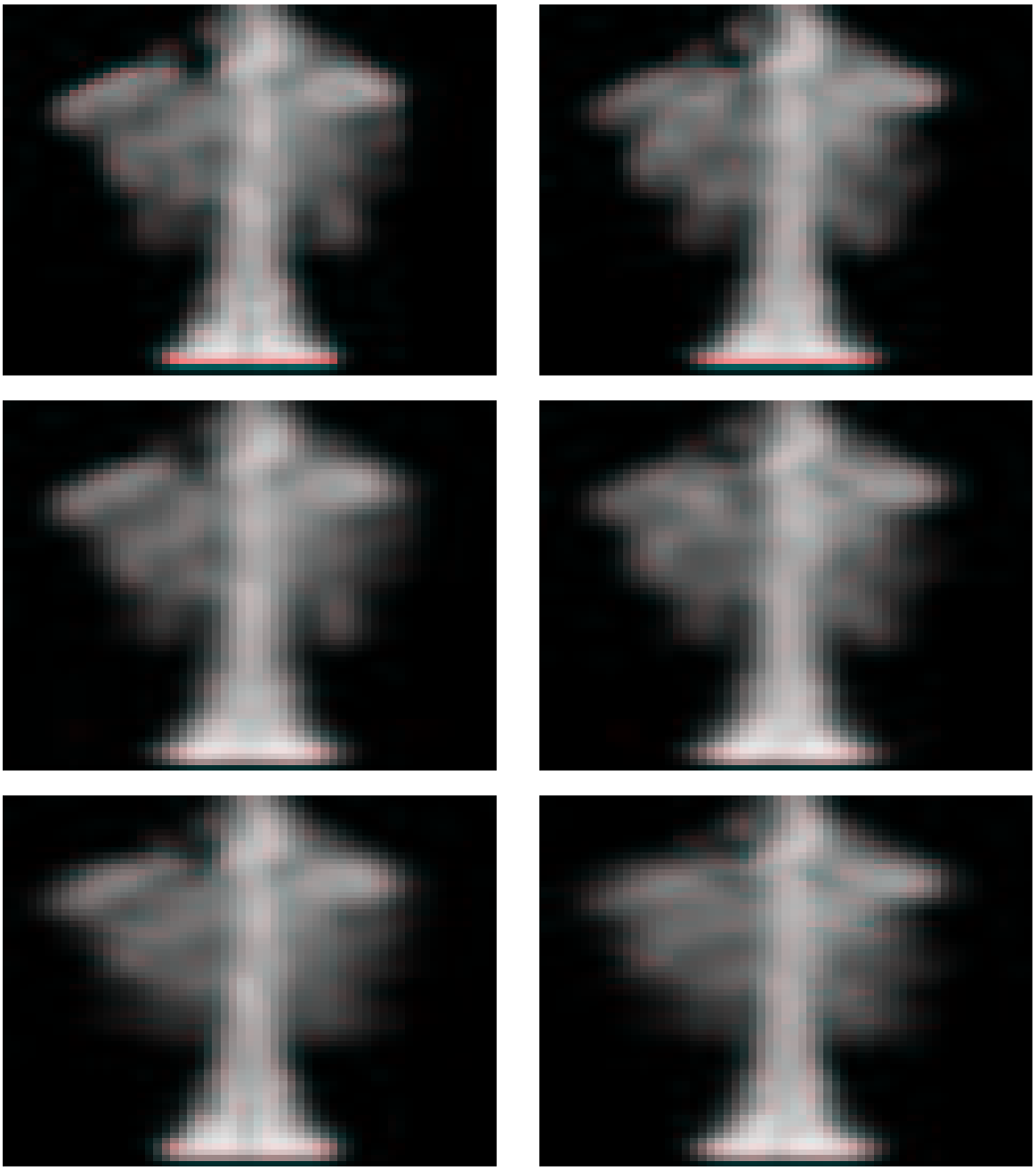


Figure 5.18: Reconstruction of the HOOD dataset using 5 input images. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. The left column corresponds to an input view and the right column to a novel intermediate view.

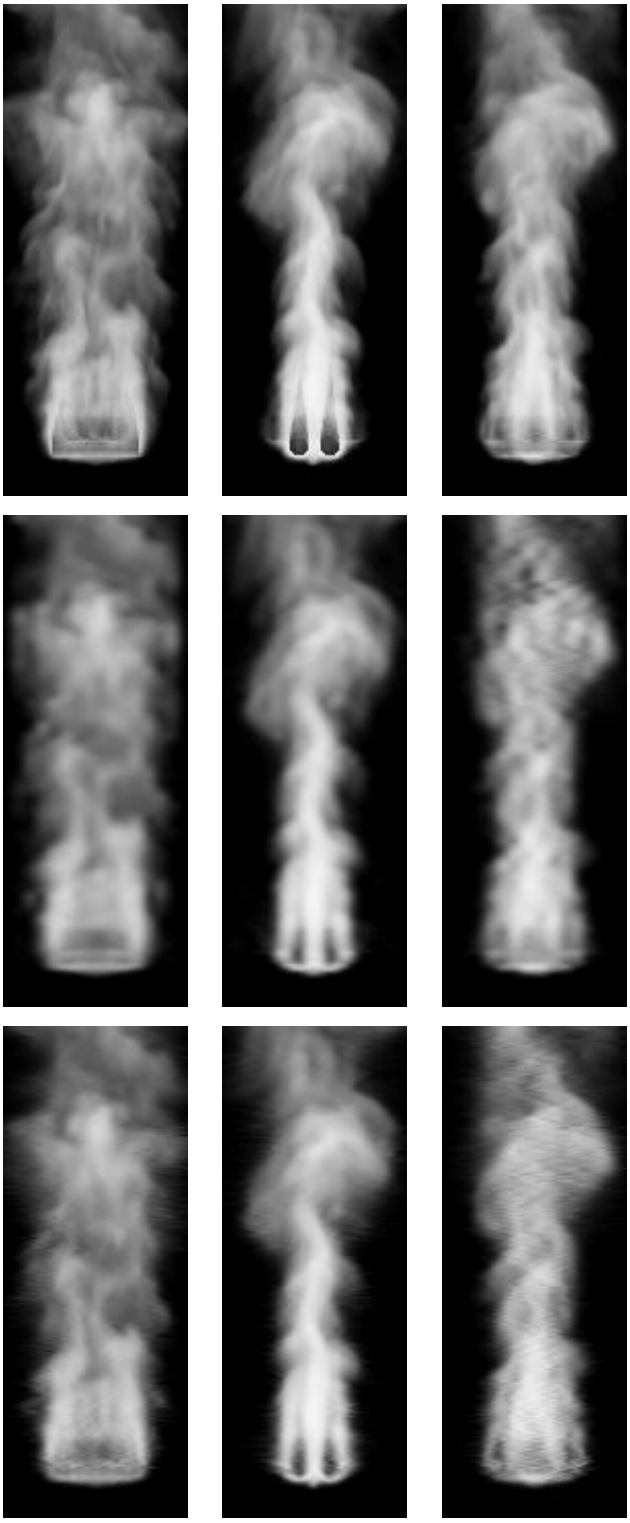


Figure 5.19: Reconstruction of the CAMPFIRE dataset using 5 input images. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. The left and middle columns correspond to input views and the right column corresponds to a novel intermediate view.

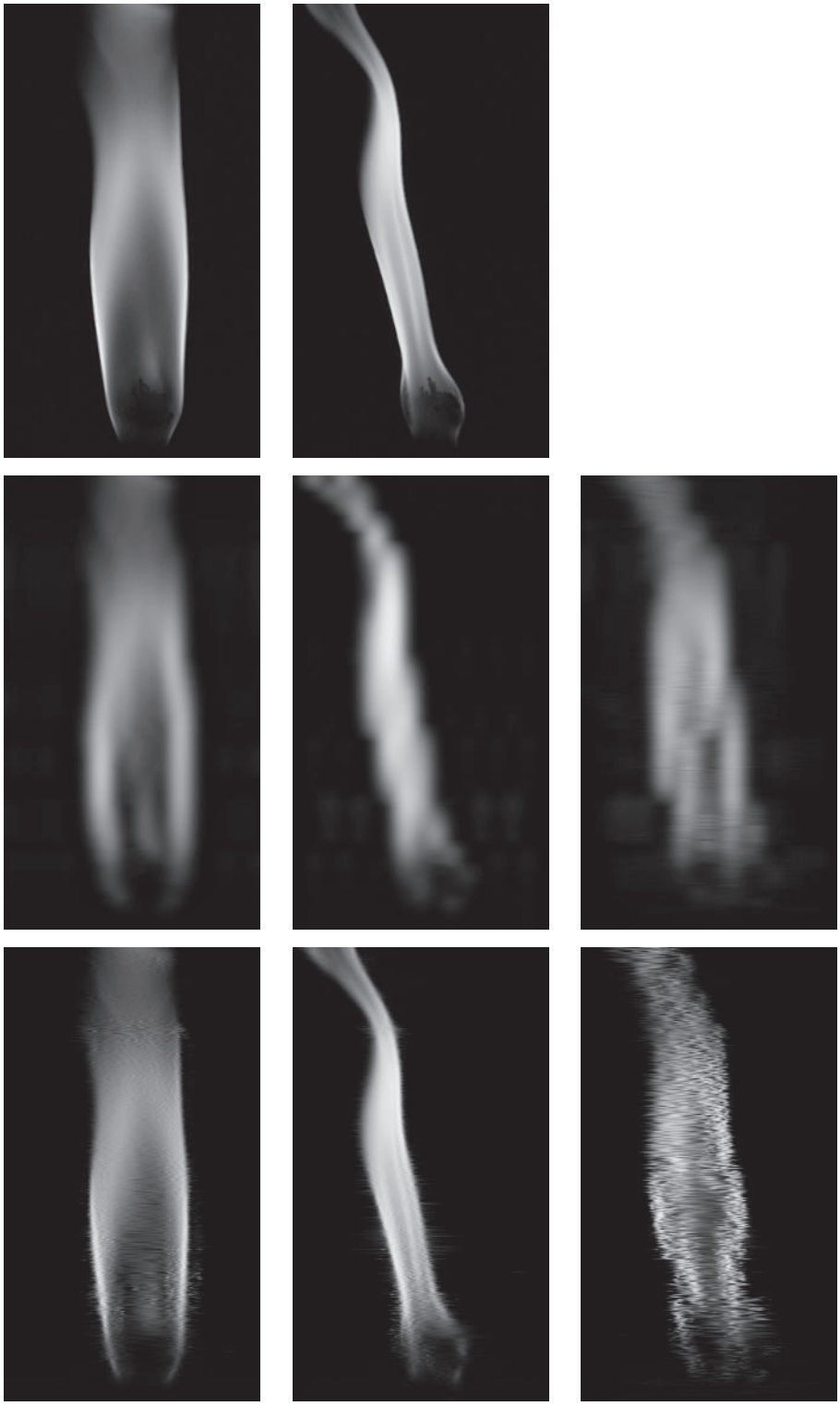


Figure 5.20: Reconstruction of the TORCH dataset using 2 input images. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. The left and middle columns correspond to the input views and the right column corresponds to a novel intermediate view.

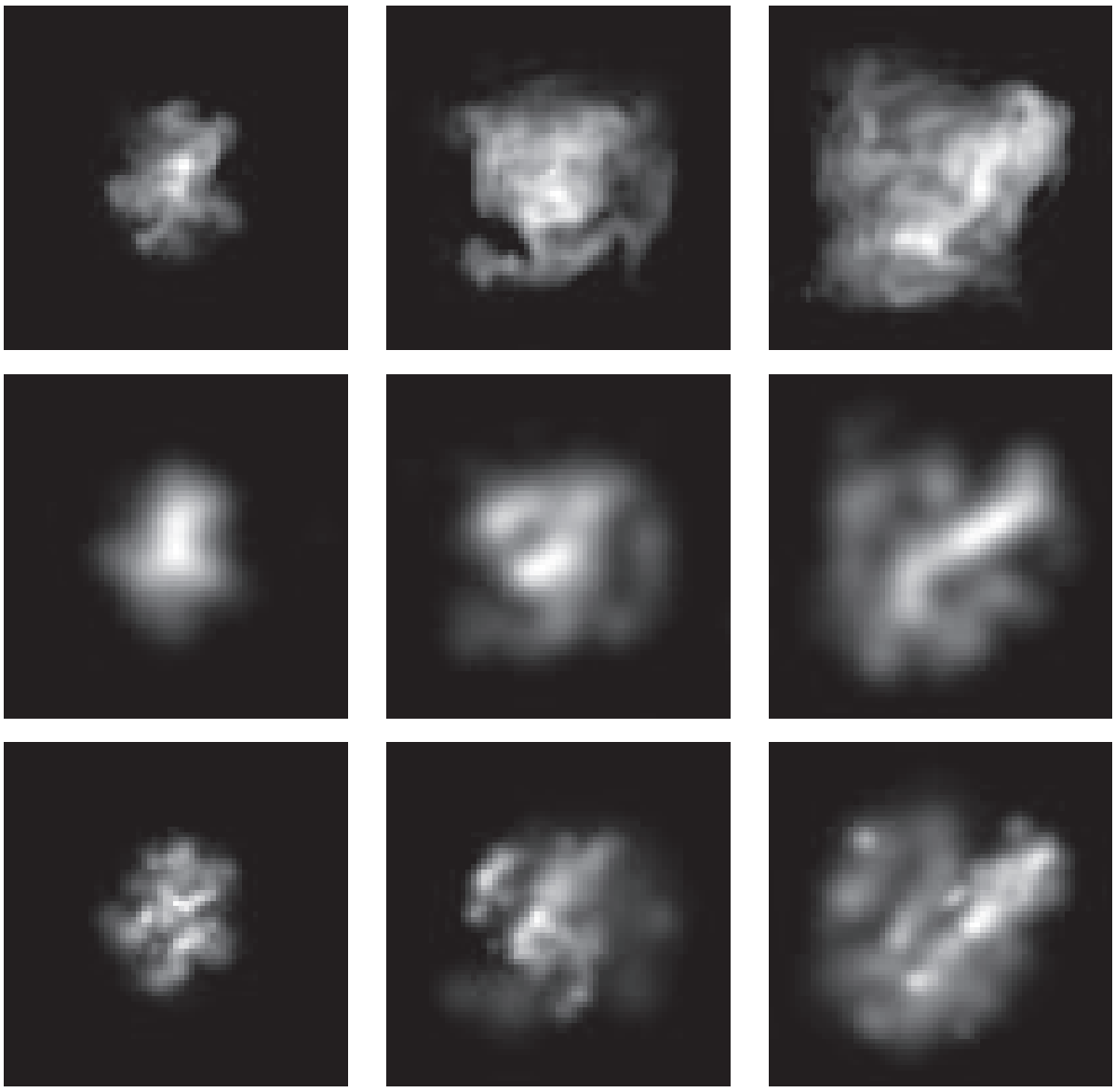


Figure 5.21: Comparing the reconstructed fire density (using 5 input images) and the ground truth for slices of the PLUME dataset. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. Each column corresponds to a different parallel slice of the reconstruction.

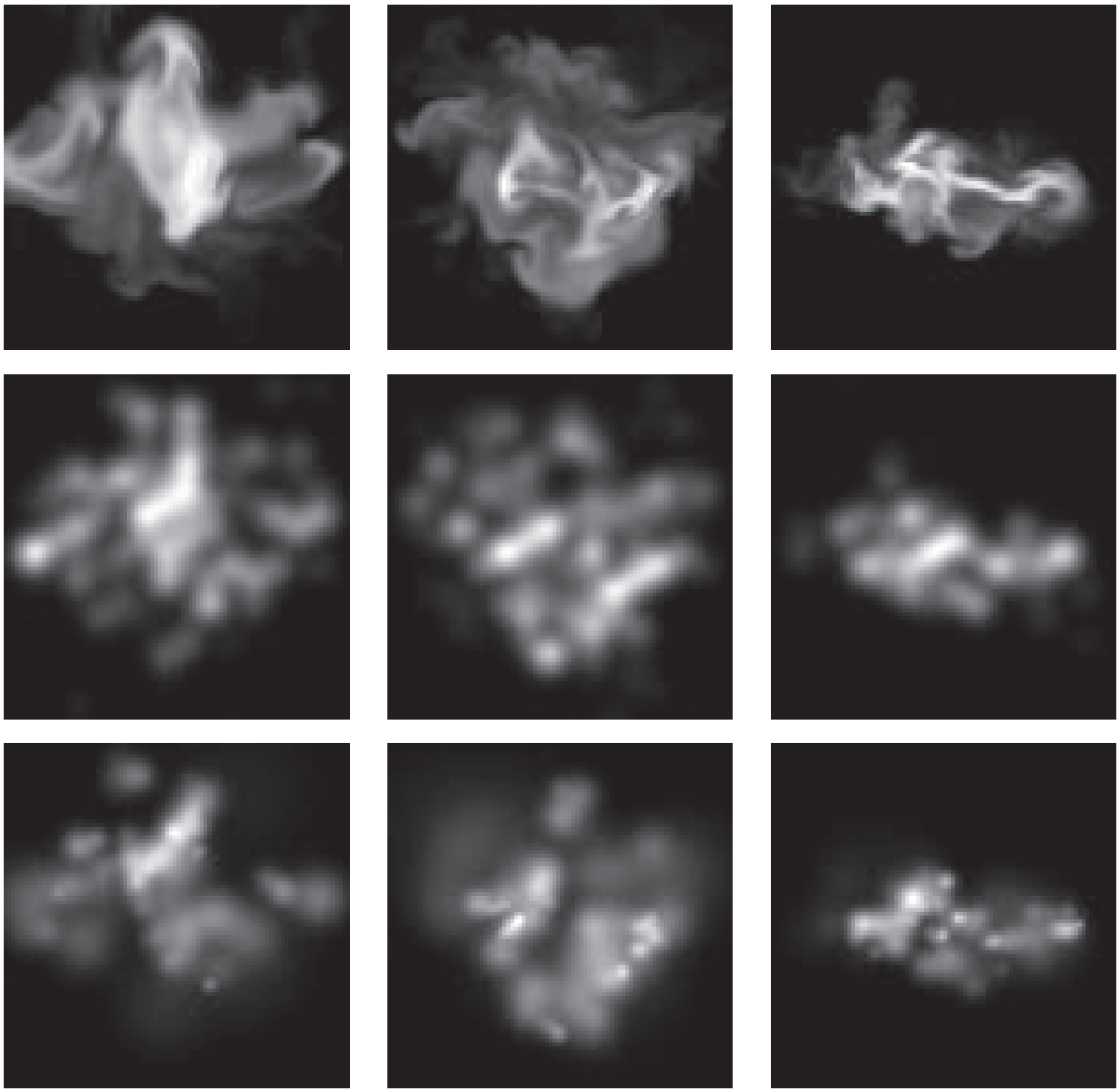


Figure 5.22: Comparing the reconstructed fire density (using 5 input images) and the ground truth for slices of the CAMPFIRE dataset. From top to bottom: ground truth, coarse reconstruction, and final reconstruction. Each column corresponds to a different parallel slice of the reconstruction.

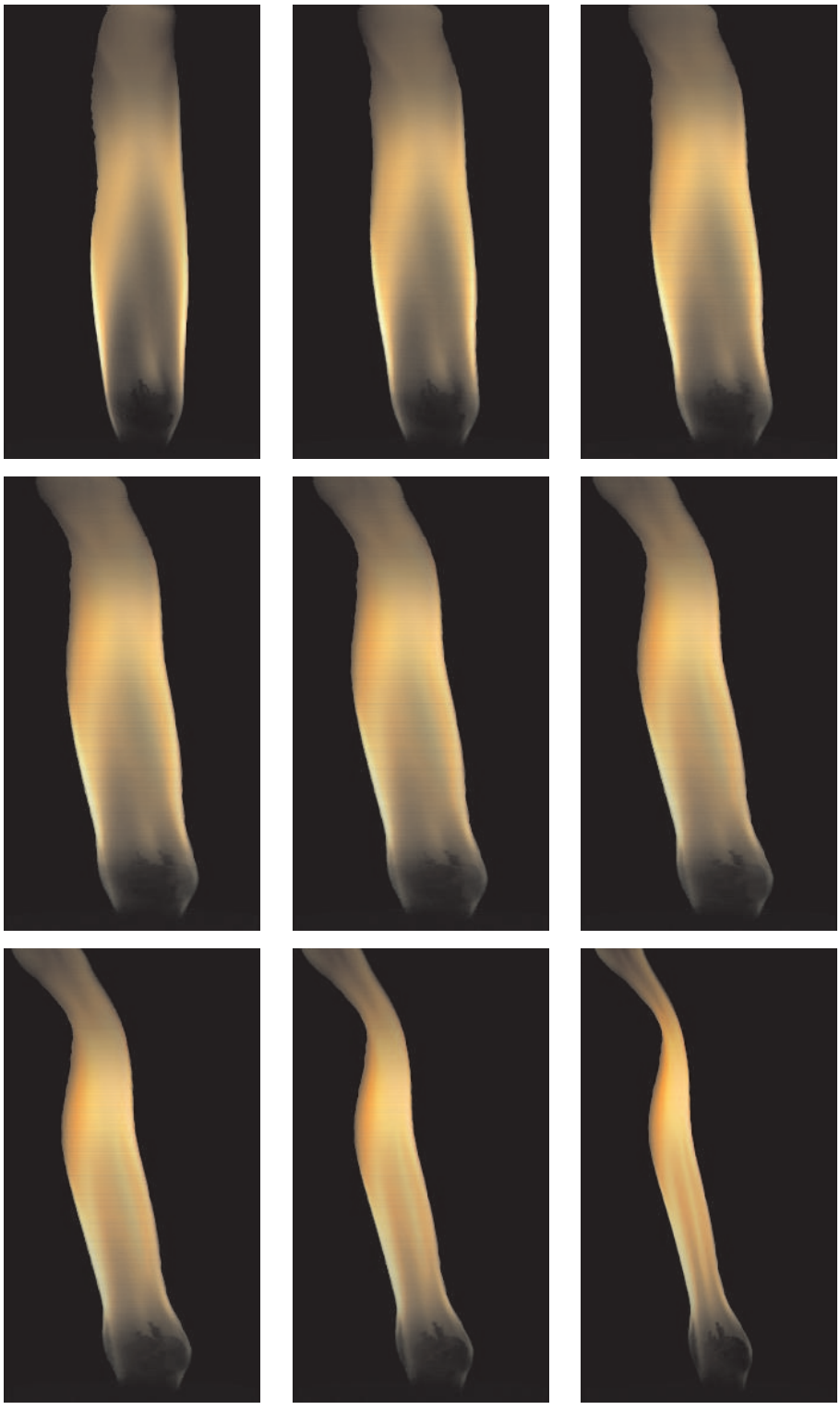


Figure 5.23: Sheet reconstruction of the TORCH dataset from 2 input images. The image sequence represents a ninety-degree rotation of the camera about the vertical axis of the fire, and the first and last images correspond to the input views.

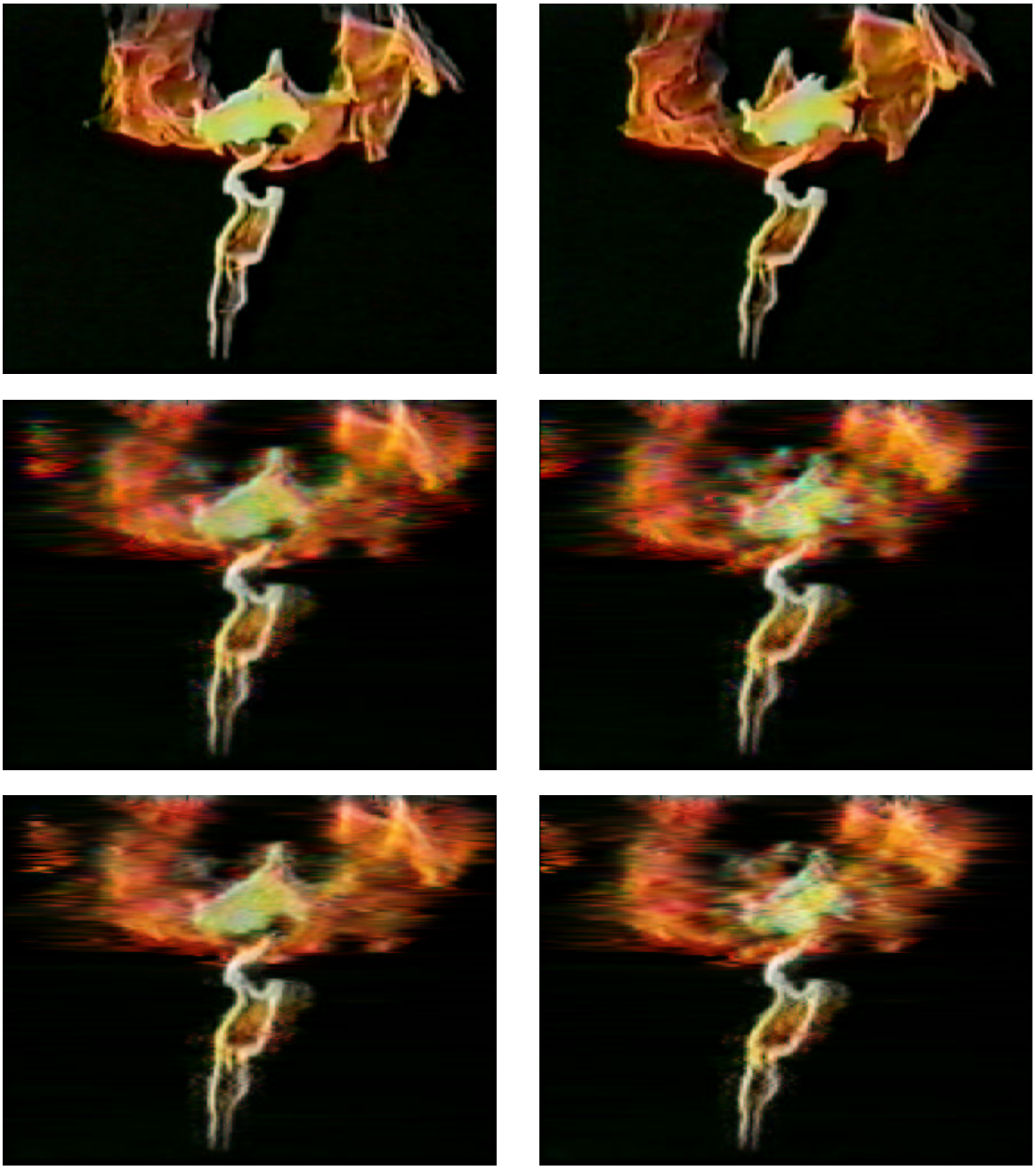


Figure 5.24: Colour reconstruction of the MATRIX dataset using 5 input images. From top to bottom: ground truth, solving the RGB channels independently, and assigning a different hue to each blob. The left column corresponds to an input view and the right column to a novel intermediate view.

the Levenberg-Marquardt method, however the high dimensionality and poor convergence behaviour present serious computational challenges.

Another open question is what are the limits on the resolution of the reconstruction in terms of the number of input images and blobs. The standard Nyquist limit does not apply here because our sampling pattern is non-uniform [19, 118]. Further complicating this analysis, we would like to integrate prior information about fire such as spatial coherence.

According to the physics of fire, extinction, refraction and other properties are all wavelength dependent. This suggests that perhaps we should deal with the RGB channels separately when performing fire reconstruction in colour. On the other hand, the RGB channels are not pure spectral colours, and separating the channels introduces colour bleeding artefacts where blobs do not align perfectly. Another idea is to use the same set of blobs for all three RGB channels, but solve for the blob weights separately. In this way each blob is assigned a hue based on the three different weights. These two colour reconstruction methods are illustrated for the MATRIX dataset with five input images in Figure 5.24, though both methods demonstrate unsatisfying artefacts. Future work will address more sophisticated methods for colour reconstruction.

5.4 Image-Based Rendering Method

The most straightforward rendering method for fire is ray tracing, applying the forward imaging model presented in Equation (5.9) to the reconstructed density field. Indeed, all of the reconstruction results presented to this point were rendered using this method.

Intuitively we should be able to do even better by making use of the original input images, especially for views close to the input views. As an extreme example, if the camera position reproduces an input view exactly, we can simply redisplay the original input image. For general views, the nearby input images can be warped according to the estimated density field and these can then be merged together. This is the basic idea behind the significant area of research known as image-based rendering (see such seminal papers as [24, 72, 103]).

Recent work has explored unifying image-based and volumetric rendering, but this work has concentrated mainly on the case where most rays reach full opacity [77, 23]. By contrast, we consider relatively thin and transparent fire, assuming that the dynamic range of imaging system is sufficient to avoid significant saturation. The thinness of

fire is both a simplification and challenge for image-based rendering. Because fire is so transparent, issues of visibility and hole-filling that arise from the image-based rendering of opaque objects can be essentially ignored. On the other hand, the lack of opaque surfaces makes it more difficult to localize image intensity in the volume for the purposes of warping.

We first present a variation on a technique due to Mueller *et al.* [77] for image-based volumetric rendering. This technique uses brute-force to distribute pixel intensity over a large number of layers, giving in a sense the best quality that could be expected from such an image-based rendering method. We then present a novel technique based on layered depth images [103] which achieves comparable results at much higher speeds.

5.4.1 Many-Layered Volume Rendering

One straightforward approach to the image-based volume rendering of fire involves uniformly decomposing the volume, for each input view, into a large number of transparent layers perpendicular to the viewing direction [77]. This method, we call the many-layered volume rendering (MLVR) method, is illustrated in Figure 5.25. We choose the number of layers according to the number of pixels in the 1D input images.

For each image pixel in a particular input view, we backproject its intensity onto the volume, to the points where the corresponding ray through the density field intersects all the different layers. We distribute this intensity proportionally over the intersection points, according to the density of the reconstructed fire over the different points along the ray. Each intersection point thus approximates a partial line integral of the density field "corrected" by the true input pixel from this view.

Once the backprojection step has been performed for all input views, rendering then consists of warping the redistributed pixel intensities onto novel rays. Because a novel ray will not necessarily meet the intersection points for some input view exactly, we use cubic interpolation within each layer to obtain estimates for the corresponding partial line integrals. By summing these partial line integrals for a novel ray, we obtain the input image warped to that ray according the density field.

Given a novel view we wish to render, we select the neighbouring input views according to the Delaunay triangulation of the camera positions. These input views are warped to the novel viewpoint as previously described, and are combined linearly according to weights based on the inverse distance to the novel view.

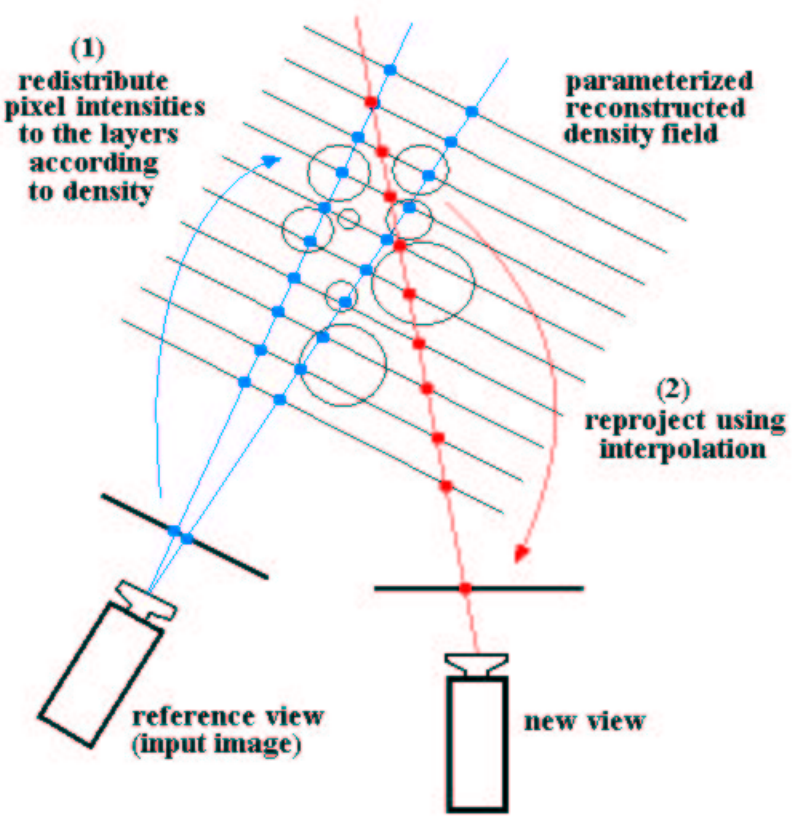


Figure 5.25: Illustration of the many-layered volume rendering (MLVR) method for the image-based volume rendering of fire.

5.4.2 Layered Depth Images for Volume Rendering

Another approach to the image-based rendering of fire involves extending the framework for layered depth images (LDIs) [103] to volume rendering. In the regular LDI framework, every pixel is associated with some small number of billboards (layers) along the corresponding ray, where the ray intersects different opaque surfaces in the scene. For highly concave opaque objects close the camera, the multiple depths associated with each pixel improves the treatment of visibility and reduces common image-based rendering artefacts associated with parallax.

In order to extend the LDI framework to the volumetric case we must determine appropriate locations to place the transparent billboards according to the reconstructed density field. A straightforward and effective solution is to associate the billboards with those blobs along the viewing ray that contribute the most to the pixel intensity. As illustrated in Figure 5.26, we define the locations of the billboards based on the projections of the centres of the most significant blobs onto the viewing ray. We then backproject pixel

intensity onto the the billboards according to the fire density as in the MLVR method.

For rendering, as in the standard LDI framework, splatting is used to accumulate the footprints of various transparent billboards to obtain the input image warped according to some novel view. As before, the different warped input images can be combined based on how close the input views are to the novel view.

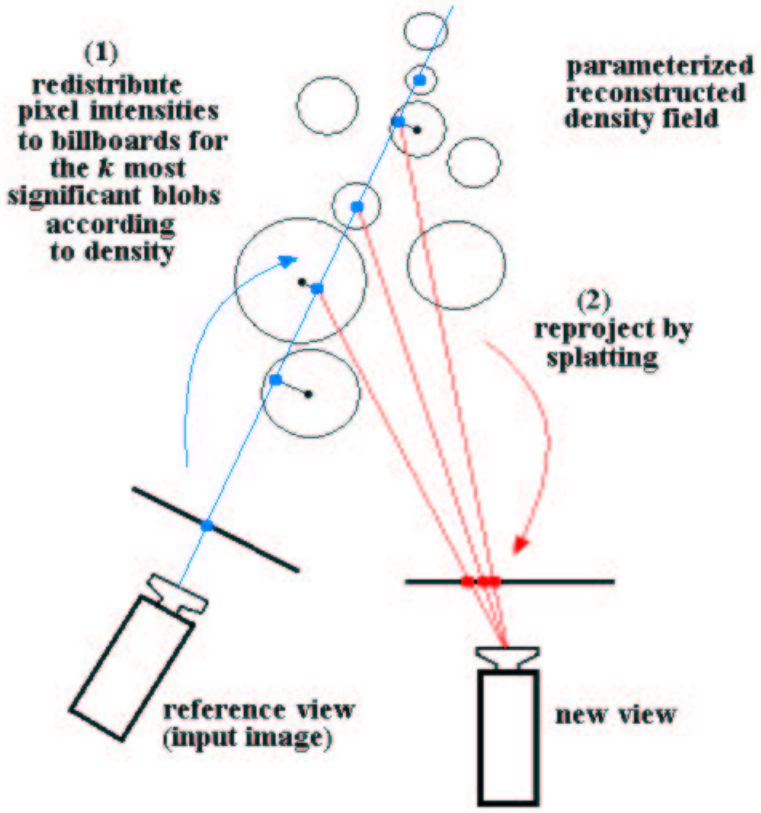


Figure 5.26: Illustration of layered depth images (LDIs) for the image-based volume rendering of fire.

5.4.3 Results and Discussion

As shown in Table 5.3 for the MATRIX dataset with five input images, both of the image-based volumetric rendering techniques constitute a significant numerical improvement over the standard ray tracing forward projection method. Note that the error for the input images is zero, as it should be, for all of the image-based rendering methods.

While the MLVR method with a very high number of layers in a sense represents the best quality we could possibly expect from such an image-based volumetric rendering method, the volumetric LDI method with 10 or more layers achieves similar numerical

rendering technique	MSE (input)	MSE (all)
ray tracing	271.5	330.7
LDI(3)	0	234.9
LDI(5)	0	221.7
LDI(10)	0	216.2
LDI(25)	0	215.2
LDI(100)	0	215.2
MLVR	0	215.1

Table 5.3: Comparing the mean squared error of reconstructions using different rendering methods for the MATRIX dataset with 5 input images. The MLVR method used the same number of layers as there were pixels in the input images (159). The number of layers used by the volumetric LDI method is indicated in brackets.

results. Assuming that each depth layer requires a separate pass with typical graphics hardware, having as many as 10 such layers may prohibit truly interactive frame rates. However, the volumetric LDI method is still much faster than the MLVR method.

If very few layers are used for the volumetric LDI approach, depth discontinuities between adjacent pixels in layers of the LDI may cause rendering artefacts. Using a larger number of layers, in the spirit of the MLVR approach, reduces the efficiency of rendering but also seems to subsume the need to worry about maintaining coherence. We have considered explicitly penalizing the lack of coherence between layer depths for adjacent LDI pixels, perhaps by using a graph-cut formulation to approximate the optimal solution [94, 17]. However, this only seems worth the effort when the number of LDI layers needs to be severely restricted.

The different rendering methods are compared for the MATRIX dataset in Figure 5.27, and the corresponding difference images are shown in Figure 5.28. As shown, the LDI and MLVR image-based methods give comparable results for rendering a static intermediate frame, and both of these methods give somewhat better results than ray tracing.

Qualitatively, the image-based rendering results are somewhat less satisfying that the significant improvements in mean squared error would lead us to believe. Errors in the reconstructed volume (partially due to poor camera calibration for the MATRIX dataset) lead to significant ghosting artefacts when the warped images of nearby input views are combined. We can see this problem in Figure 5.27, where ambiguity in the reconstruction causes a connected triangular shape to be split into three pieces in the intermediate views. While this sort of ambiguity artefact appears in basic ray traced renderings as well, the image-based rendering approach makes this artefact more noticeable when the fire is

viewed in animation with a smoothly varying viewpoint. In animation, the input views appear to fade in and out unnaturally in the regions where the shape of the fire is poorly reconstructed.

For the synthetic HOOD dataset, where calibration is not an issue, the image-based rendering technique seems to perform qualitatively better. On the other hand, the thick plume of fire and the relatively low resolution make this somewhat difficult to assess. In Table 5.4, we compare the effect of using less layers for the MLVR method, and illustrate these results in Figure 5.29. Corresponding difference images are shown in Figure 5.30. Note that restricting the number of MLVR layers produces speckle artefacts, and that a large number of layers (half the resolution of the input images) is required to produce acceptable results.

rendering technique	MSE (input)	MSE (all)
ray tracing	163.0	169.4
MLVR(10)	0	22.7
MLVR(30)	0	16.0
MLVR(59)	0	15.7

Table 5.4: Comparing the mean squared error of reconstructions using different numbers of layers for MLVR rendering method for the HOOD dataset with 5 input images. The number of layers used by the image-based MLVR method is indicated in brackets. The number of pixels in the input images is 59.

5.5 Summary

We have addressed the reconstruction of fire from synchronized images using a physically-based imaging model. Under our assumed imaging model, the problem was shown to be equivalent to computerized tomography with very few input images. Our analysis demonstrated that the space of solutions that project to the input images is fundamentally ambiguous, and two extreme analytic solutions were presented for the two-camera case.

Next, we presented an algebraic reconstruction method for fire that exploits the parameterization in terms of blobs to reduce the dimensionality of the problem and implicitly ensure a spatially coherent reconstruction. By fixing some of the blob parameters, we obtain a constrained linear least-squares problem that can be solved to get a coarse initial reconstruction. We then redistribute the blobs using stochastic resampling according to

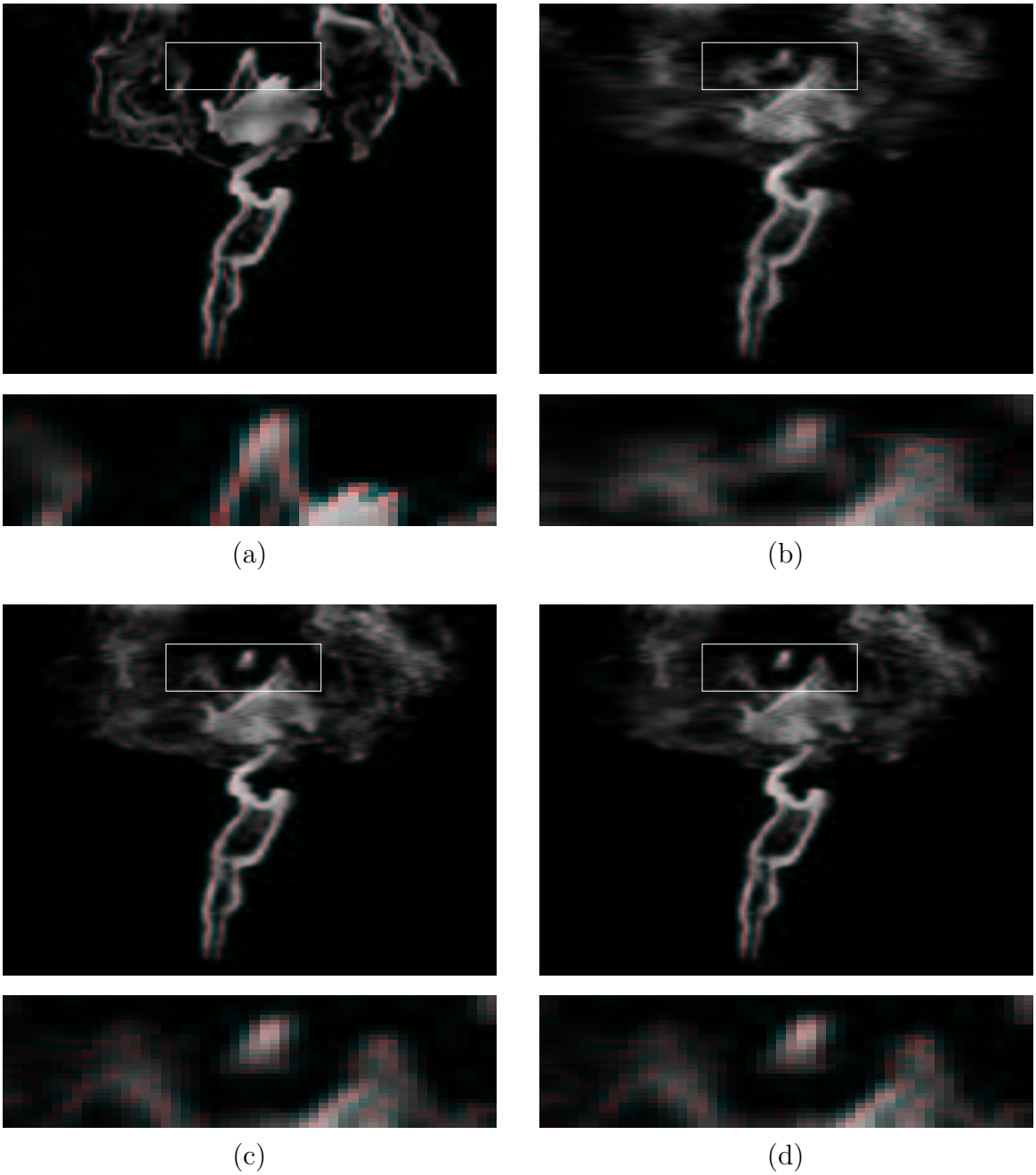


Figure 5.27: Different rendering methods applied to an intermediate view of the MATRIX dataset using 5 input images. (a) ground truth, (b) standard raytracing forward projection, (c) volumetric LDI method (10 layers), and (d) image-based MLVR method (159 layers). The white rectangles indicate a region, expanded below each image, where ambiguity in the reconstruction causes rendering artefacts.

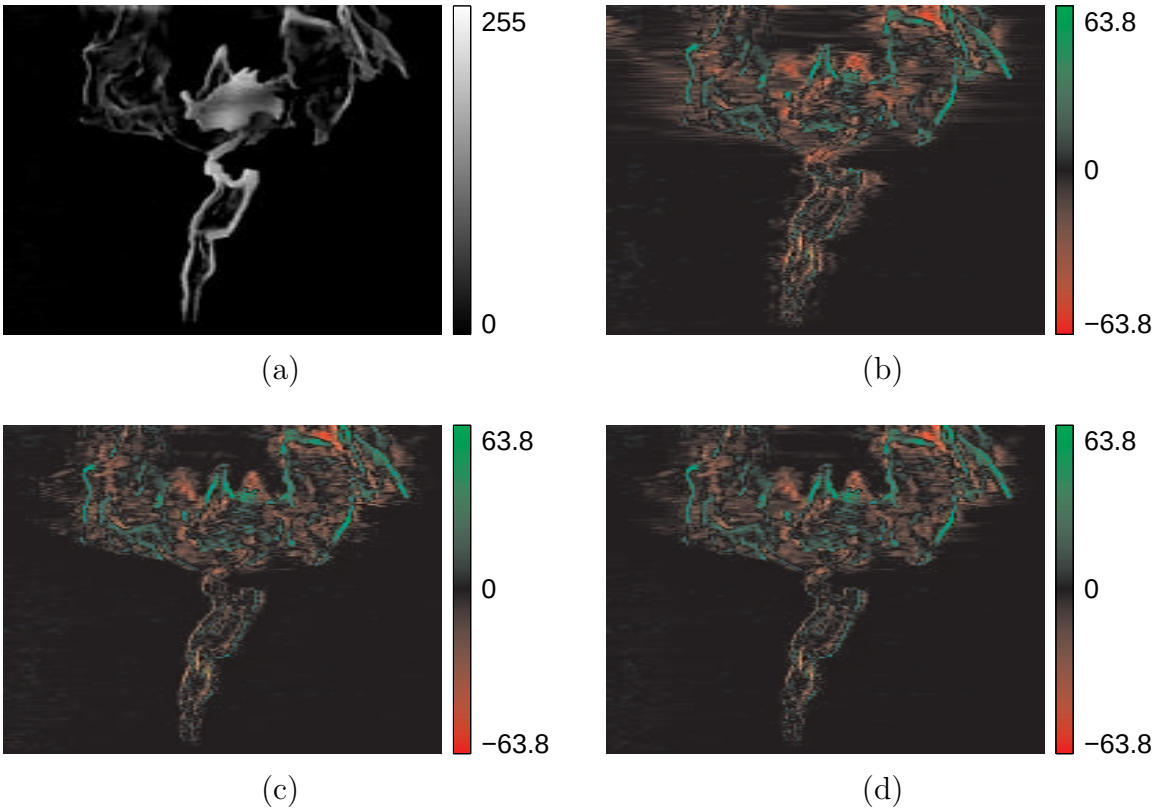
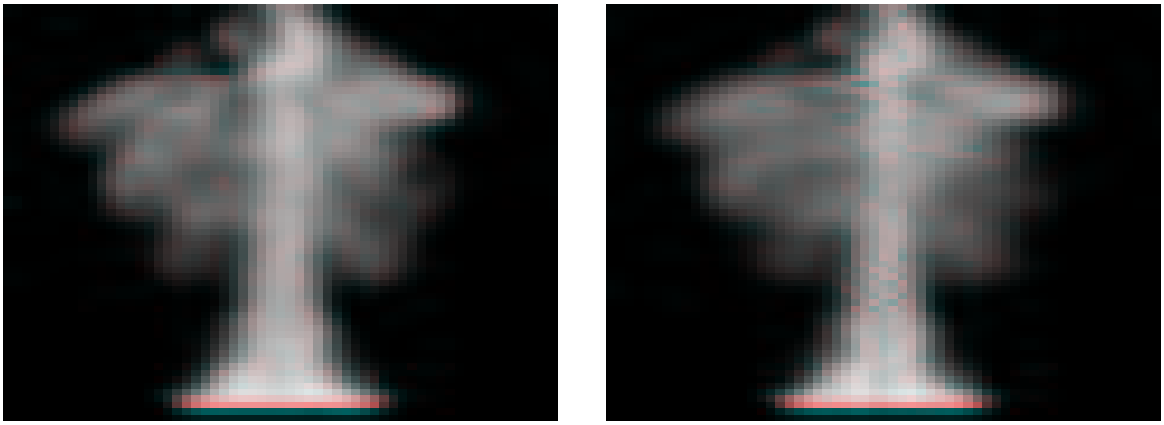
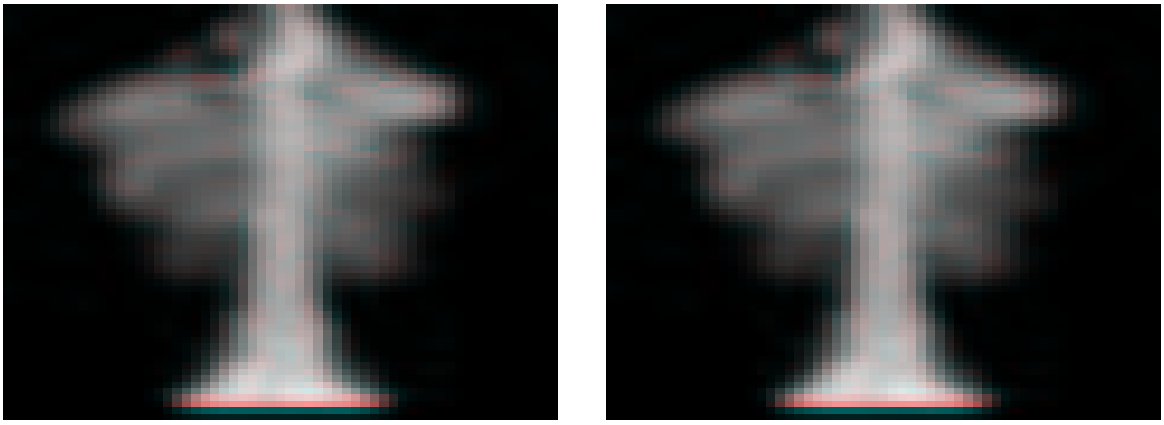


Figure 5.28: Difference images for different rendering methods applied to an intermediate view of the MATRIX dataset using 5 input images. (a) ground truth, and difference images for (b) standard raytracing forward projection, (c) volumetric LDI method (10 layers), and (d) image-based MLVR method (159 layers). The difference images are shown in false colour.



(a) (b)



(c) (d)

Figure 5.29: Modifying the number of layers for the MLVR method. An intermediate view is shown for the HOOD dataset using 5 input images. (a) ground truth, and (b)-(d) image-based MLVR method with 10, 30, and 59 layers.

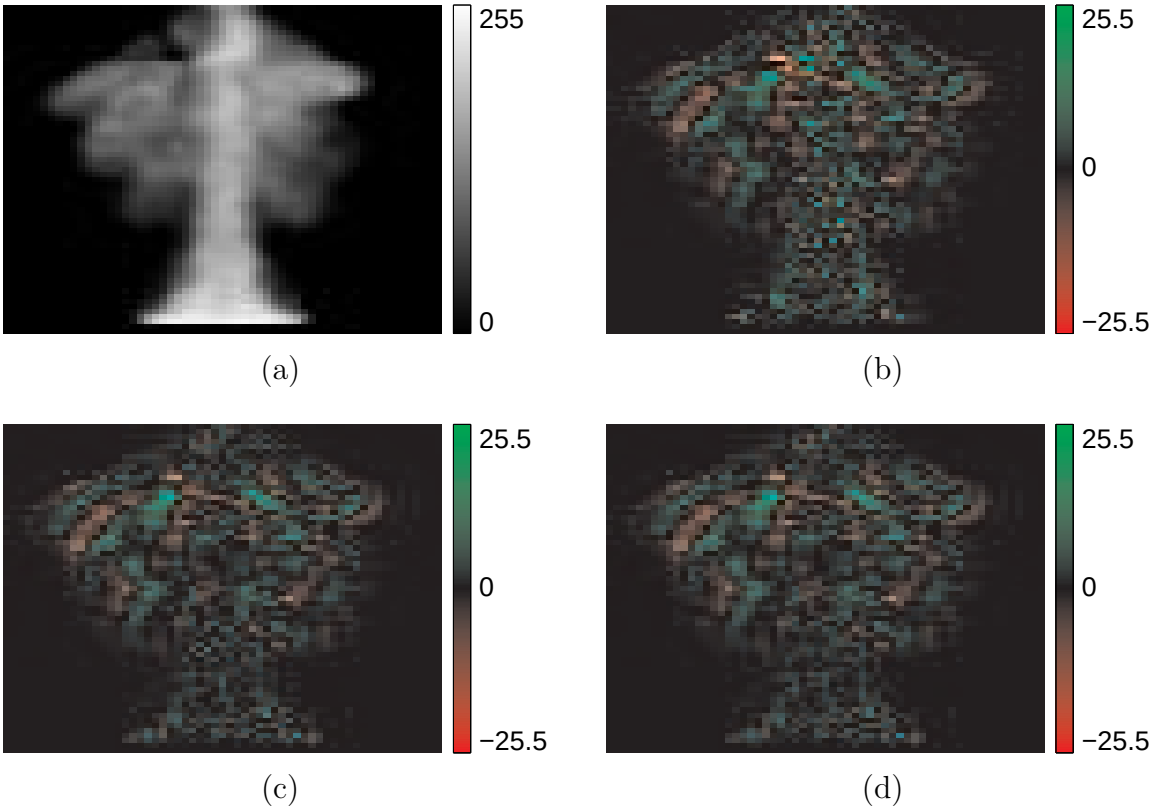


Figure 5.30: Difference images for the MLVR rendering method with different numbers of layers. An intermediate view is shown for the HOOD dataset using 5 input images. (a) ground truth, and (b)-(d) difference images for the image-based MLVR method with 10, 30, and 59 layers. The difference images are shown in false colour.

the results of the coarse reconstruction, and perform the reconstruction again to capture finer geometric detail.

To render the reconstructed fire, one possibility is to simply apply the forward projection model and ray trace. We also presented two image-based volumetric rendering schemes that exploit the information contained in the original input images. These image-based methods perform well close to the input views, but further away from the input views, ghosting artefacts are problematic in regions where the reconstruction is poor.

Future directions include using the formal analysis of ambiguity to devise new reconstruction methods for navigating the space of photo-consistent reconstructions directly, expanding the reconstruction method to recover various physical properties of fire as well, and designing an objective function that evaluates the plausibility of a given reconstruction as fire in terms of these physical properties.

Chapter 6

Conclusions

This thesis investigated the problem of 3D reconstruction of fire from synchronized images or video sequences. We formalized the problem in terms of the underlying physics of fire, presented an analysis of the space of photo-consistent solutions, and developed new methods for the image-based reconstruction and rendering of fire. In this way, we have made steps towards expanding the class of phenomena that can be reconstructed from images.

One of the key ideas to emerge from this work is that a good algebraic reconstruction depends on an appropriate parameterization of the density field, and that an iterative resampling process can greatly improve this parameterization based on previous reconstruction results. Another important idea is that the space of reconstructions is fundamentally ambiguous, which is especially significant for a sparse number of input views.

Reconstructing natural phenomena is challenging because the imaging model must account for more sophisticated effects such as volumetric transparency. However, we believe that image-based reconstruction of natural phenomena from the real world will ultimately provide greater realism than pure physical simulation. In a sense, these two approaches are complementary, because image-based reconstructions could be used to validate or train physical simulators, and the same physical models used for simulation are presumably used to drive the reconstruction as well.

For future work, we would like to expand these ideas to obtain image-based reconstructions of other natural phenomena such as smoke, haze, weather effects, and fluids. We would also like to exploit the dynamics of the various phenomena for disambiguation, in order to direct the reconstruction to the more physically plausible solutions.

Appendix A

Space of Two-Camera Reconstructions

Let $\mathcal{D}$ denote the set of $n \times n$ matrices s.t. $\mathbf{1}^T D = D\mathbf{1} = 0$ where $D \in \mathcal{D}$ and $\mathbf{1} = (1, 1, \ldots, 1)^T$. We will now prove that $\mathcal{D}$ is a vector space. Since $\mathcal{D}$ is a subset of the $n \times n$ real matrices, $M_{n\times n}(\Re)$, we need only prove that $kx+y \in \mathcal{D}$ for all $x, y \in \mathcal{D}$, $k \in \Re$. Trivially $\mathbf{1}^T(kx + y) = \mathbf{1}^T(kx) + \mathbf{1}^T y = k\mathbf{1}^T x + \mathbf{1}^T y = 0$ and similarly $(kx + y)\mathbf{1} = 0$. So $\mathcal{D}$ is in fact a vector space.

Now we establish a linear isomorphism $T : M_{(n-1)\times(n-1)}(\Re) \mapsto \mathcal{D}$. Define T as follows. Given $Y \in M_{(n-1)\times(n-1)}(\Re)$ let $a_{in} = -\sum_{j=1}^{n-1} y_{ij}$ and $a_{ni} = -\sum_{j=1}^{n-1} y_{ji}$ for i between 1 and $(n-1)$, where y_{ij} is an element of Y. In addition, let $a_{nn} = -\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} y_{ij}$. Then,

$$T(Y) = \begin{pmatrix} y_{11} & \cdots & y_{1(n-1)} & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ y_{(n-1)1} & \cdots & y_{(n-1)(n-1)} & a_{(n-1)n} \\ a_{n1} & \cdots & a_{n(n-1)} & a_{nn} \end{pmatrix}.$$

Clearly, $T(kX + Y) = kT(X) + T(Y)$ where $X, Y \in M_{(n-1)\times(n-1)}(\Re)$ and $k \in \Re$, so T is linear.

We show that T is one-to-one and onto. T is onto since $\forall D \in \mathcal{D}$, $\exists Y \in M_{(n-1)\times(n-1)}(\Re)$, namely the restriction of D to the upper-left $(n-1)\times(n-1)$ block matrix. Also a kernel of T is $\{\mathbf{0}_{(n-1)\times(n-1)}\}$ since $Y = \mathbf{0}_{(n-1)\times(n-1)}$ if $T(Y) = \mathbf{0}_{n\times n}$ by construction, thus T is one-to-one. Therefore T is an onto and one-to-one linear mapping, and thus is a linear isomorphism.

Hence, $\dim(\mathcal{D}) = (n-1)^2$ since $\dim(M_{(n-1)\times(n-1)}) = (n-1)^2$ and dimension is

preserved by a linear isomorphism.

Consider the elementary switching matrices $D_{ij} \in \mathcal{D}$, defined for i, j from 1 to $(n-1)$,

$$D_{ij} = \begin{pmatrix} +1 & \cdots & -1 & 0 \\ \vdots & 0 & \vdots & \vdots \\ -1 & \cdots & +1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

The switching matrices have four non-zero entries arranged in a rectangle, where the bottom-right $+1$ is located at row $(i+1)$ and column $(j+1)$. Now D_{ij} are linearly independent, and there are $(n-1)^2$ of them. Therefore $\{D_{ij}\}$ is a basis for $\mathcal{D}$.

For two orthogonal cameras using a parallel projection model, the space of reconstructions can be represented as $S + D$, where S is any particular solution (for example, the multiplication solution S_{mult}) and $D \in \mathcal{D}$ preserves row and colum sums. We also require that $S + D$ is a physically meaningful density, and thus is non-negative everywhere. So, the space of reconstructions is an $(n-1)^2$-dimensional hyperplane lying in $M_{n\times n}$, restricted to the region where no elements are negative, and spanned by $(n-1)^2$ elementary switching matrices. ∎

Bibliography

[1] E. Adelson and P. Anandan. Ordinal characteristics of transparency. In *Proc. AAAI Workshop on Qualitative Vision 1990*, pages 77–81, 1990.

[2] E. Adelson and A. Pentland. The perception of shading and reflectance. In D. Knill and W. Richards, editors, *Perception as Bayesian Inference*, pages 409–423, New York, 1996. Cambridge University Press.

[3] A. K. Agrawal, N. Butuk, S. R. Gollahalli, and D. Griffin. Three-dimensional rainbow schlieren tomography of a temperature field in gas flows. *Applied Optics*, 37(3):479–485, 1998.

[4] B. W. Albers and A. K. Agrawal. Schlieren analysis of an oscillating gas-jet diffusion flame. *Combustion and Flame*, 199:84–94, 1999.

[5] A. Andersen and A. Kak. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6:81–94, 1984.

[6] D. Arquès, E. Felgines, S. Michelin, and K. Zampieri. Thermal convection in turbulent flow, image synthesis and heat animation. In *Proc. International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media 1999*, pages 337–344, 1999.

[7] S. Baker, R. Szeliski, and P. Anandan. A layered approach to stereo reconstruction. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 1998*, pages 434–441, 1998.

[8] R. T. Baum, K. B. McGratten, and M. Nyden. An examination of the applicability of computed tomography for the measurement of component concentrations in fire-generated plumes. *Combustion and Flame*, 113:358–372, 1998.

[9] P. Beaudoin, S. Paquet, and P. Poulin. Realistic and controllable fire simulation. In *Proc. Graphics Interface 2001*, pages 159–166, 2001.

[10] J. Beck and R. Ivry. On the role of figural organization in perceptual transparency. *Perception and Psychophysics*, 44(6):585–594, 1988.

[11] M. Berger, T. Trout, and N. Levit. Ray tracing mirages. *IEEE Computer Graphics and Applications*, 10(3):36–41, May–June 1990.

[12] M. Bhatia, W. Karl, and A. Willsky. A wavelet-based method for multiscale tomographic reconstruction. *IEEE Transactions on Medical Imaging*, 15(1):92–101, 1996.

[13] R. Bhotika, D. J. Fleet, and K. N. Kutulakos. A probabilistic theory of occupancy and emptiness. In *Proc. European Conference on Computer Vision 2002*, volume 3, pages 112–132, 2002.

[14] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998.

[15] J. D. Bonet and P. Viola. Roxels: Responsibility weighted 3D volume reconstruction. In *Proc. IEEE International Conference On Computer Vision 1999*, pages 418–425, 1999.

[16] J.-Y. Bouguet. Camera Calibration Toolbox for Matlab, 06-September-2002. http://www.vision.caltech.edu/bouguetj/calib_doc/.

[17] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. In *Proc. IEEE International Conference On Computer Vision 1999*, pages 377–384, 1999.

[18] C. K. Brain and A. Sillen. Evidence from the Swartkrans cave for the earliest use of fire. *Nature*, (336):464–466, Dec. 1988.

[19] Y. Bresler and P. Feng. Spectrum-blind minimum-rate sampling and reconstruction of 2-D multiband signals. In *Proc. IEEE International Conference on Image Processing 1996*, pages 701–704, 1996.

[20] Y. Bresler, J. A. Fessler, and A. Macovski. A Bayesian approach to reconstruction from incomplete projections of a multiple object 3D domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):840–858, Aug. 1989.

[21] A. Broadhurst, T. Drummond, and R. Cipolla. A probabilistic framework for the Space Carving algorithm. In *Proc. IEEE International Conference of Computer Vision 2001*, pages 388–393, 2001.

[22] R. W. Bukowski and C. H. Séquin. Interactive simulation of fire in virtual building environments. In *Proc. ACM SIGGRAPH 1997*, pages 35–44, 1997.

[23] B. Chen, A. Kaufman, and Q. Tang. Image-based volume rendering. In *Proc. IEEE Workshop on Volume Graphics 2001*, 2001.

[24] S. E. Chen and L. Williams. View interpolation for image synthesis. In *Proc. ACM SIGGRAPH 1993*, pages 279–288, 1993.

[25] N. Chiba, K. Muraoka, H. Takahashi, and M. Miura. Two-dimensional visual simulation of flames, smoke and the spread of fire. *The Journal of Visualization and Computer Animation*, 5(1):37–54, Jan.–Mar. 1994.

[26] Y. Chuang, B. Curless, D. Salesin, and R. Szeliski. A Bayesian approach to digital matting. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 2001*, pages 264–271, 2001.

[27] Y. Chuang, D. E. Zongker, J. Hindorff, B. Curless, D. H. Salesin, and R. Szeliski. Environment matting extensions: Towards higher accuracy and real-time capture. In *Proc. ACM SIGGRAPH 2000*, pages 121–130, 2000.

[28] Y.-Y. Chuang, A. Agarwala, B. Curless, D. Salesin, and R. Szeliski. Video matting of complex scenes. In *Proc. ACM SIGGRAPH 2002*, pages 243–248, 2002.

[29] F. Cignoli, S. D. Iuliis, V. Manta, and G. Zizak. Two-dimensional two-wavelength emission technique for soot dignostics. *Applied Optics*, 41(30):5370–5378, 2001.

[30] R. T. Collins. A space-sweep approach to true multi-image matching. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 1996*, pages 358–363, 1996.

[31] D. P. Correia, P. Ferrao, and A. Caldeira-Pires. Advanced 3D emission tomography flame temperature sensor. *Combustion Science and Technology*, 163:1–24, 2001.

[32] Digital Air Incorporated. http://www.virtualcamera.com.

[33] Y. Dobashi, K. Kaneda, H. Yamashita, T. Okita, and T. Nishita. A simple, efficient method for realistic animation of clouds. In *Proc. ACM SIGGRAPH 2000*, pages 19–28, 2000.

[34] R. A. Drebin, L. Carpenter, and P. Hanrahan. Volume rendering. In *Proc. ACM SIGGRAPH 1988*, pages 65–74, 1988.

[35] A. C. Drescher, A. J. Gadgil, P. N. Price, and W. W. Nazaroff. Novel approach for tomographic reconstruction of gas concentration distributions in air: Use of smooth basis functions and simulated annealing. *Atmospheric Environment*, 30(6):929–940, 1996.

[36] D. Drysdale. *An Introduction to Fire Dynamics*. John Wiley and Sons, Chichester, second edition, 1998.

[37] J. Duderstadt and W. Martin. *Transport Theory*. John Wiley and Sons, New York, NY, 1979.

[38] V. Elangovan. A direct method for segmenting tomographic data. Master's thesis, University of Utah, School of Computing, 2001.

[39] V. Elangovan and R. T. Whittaker. From sinograms to surfaces: A direct approach to the segmentation of tomographic data. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention 2001*, pages 213–223, 2001.

[40] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Y. Zeevi. The farthest point strategy for progressive image sampling. *IEEE Transactions on Image Processing*, 6(9):1305–1315, Sept. 1997.

[41] G. W. Faris and R. L. Byer. Beam-deflection optical tomography. *Optics Letters*, 12(3):72–74, 1987.

[42] G. W. Faris and R. L. Byer. Beam-deflection optical tomography of a flame. *Optics Letters*, 12(3):155–157, 1987.

[43] R. Fedkiw, J. Stam, and H. W. Jensen. Visual simulation of smoke. In *Proc. ACM SIGGRAPH 2001*, pages 15–22, 2001.

[44] M. L. Fischer, P. N. Price, T. L. Thatcher, C. A. Schwalbe, M. J. Craig, E. E. Wood, R. G. Sextro, and A. J. Gadgil. Rapid measurements and mapping of tracer gas concentrations in a large indoor space. *Atmospheric Environment*, 35(16):2837–2844, 2001.

[45] N. Foster and D. N. Metaxas. Modeling the motion of a hot, turbulent gas. In *Proc. ACM SIGGRAPH 1997*, pages 181–188, 1997.

[46] T. Frese, C. A. Bouman, and K. Sauer. Adaptive wavelet graph model for Bayesian tomographic reconstruction. *IEEE Transactions on Image Processing*, 11(7):756–770, July 2002.

[47] D. Gering and W. Wells III. Object modeling using tomography and photography. In *Proc. IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes 1999*, pages 11–18, 1999.

[48] E. J. Gibson and R. D. Walk. The visual cliff. *Scientific American*, 202(4):67–71, 1960.

[49] K. Girodias, H. Barrett, and R. Shoemaker. Parallel simulated annealing for emission tomography. *Physics in Medicine and Biology*, 36(7):921–938, 1991.

[50] R. Gordon, R. Bender, and G. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology*, 29:471–481, 1970.

[51] K. Hanson and G. Wecksung. Bayesian approach to limited-angle reconstruction in computed tomography. *Journal of the Optical Society of America*, 73(11):1501–1509, 1983.

[52] R. A. Hashamonay, M. G. Yost, and C. Wu. Computed tomography of air pollutants using radial scanning path-integrated optical remote sensing. *Atmospheric Environment*, 33(2):267–274, 1999.

[53] G. T. Herman and A. Kuba, editors. *Discrete Tomography: Foundations, Algorithms, and Applications*, Boston, 1999. Birkhäuser.

[54] V. Interrante, H. Fuchs, and S. M. Pizer. Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3(2), Apr.–June 1997.

[55] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, 12(1):5–16, Feb. 1994.

[56] F. D. IX, K. Mueller, and A. Kaufman. Volumetric backprojection. In *Proc. IEEE Symposium on Volume Visualization 2000*, pages 109–117, 2000.

[57] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Proc. ACM SIGGRAPH 2001*, pages 511–518, 2001.

[58] S. Ju, M. J. Black, and A. D. Jepson. Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 1996*, pages 307–314, 1996.

[59] J. T. Kajiya and B. P. V. Herzen. Ray tracing volume densities. In *Proc. ACM SIGGRAPH 1983*, pages 165–174, 1983.

[60] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, New York, 1988.

[61] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1987.

[62] S. A. King, R. A. Crawfis, and W. Reid. Fast volume rendering and animation of amorphous phenomena. In *Proc. International Workshop on Volume Graphics 1999*, pages 229–242, 1999.

[63] K. N. Kutulakos. Approximate N-view stereo. In *Proc. European Conference on Computer Vision 2000*, volume 1, pages 67–83, 2000.

[64] K. N. Kutulakos and S. M. Seitz. A theory of shape by shape carving. *International Journal of Computer Vision*, 38(3):197–216, July 2000.

[65] A. Lamorlette and N. Foster. Structural modeling of flames for a production environment. In *Proc. ACM SIGGRAPH 2002*, pages 729–735, 2002.

[66] M. S. Langer and S. W. Zucker. Shape-from-shading on a cloudy day. *Journal of the Optical Society of America A*, 11(2):467–478, 1994.

[67] D. Laur and P. Hanarahan. Hierarchical splatting: a progressive refinement algorithm for volume rendering. In *Proc. ACM SIGGRAPH 1991*, pages 285–288, 1991.

[68] M. Levoy. Efficient ray tracing of volume data. In *Proc. ACM SIGGRAPH 1990*, pages 245–261, 1990.

[69] R. Marabini, G. T. Herman, and J. M. Carazo. 3D reconstruction in electron microscopy using art with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy*, 72(1–2):53–65, 1998.

[70] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, June 1995.

[71] P. S. McCormick and J. P. Ahrens. Visualization of wildfire simulations. *IEEE Computer Graphics and Applications*, 18(2):17–19, Mar.–Apr. 1998.

[72] L. McMillan and G. Bishop. Plenoptic modeling: An image-based rendering system. In *Proc. ACM SIGGRAPH 1995*, pages 39–46, 1995.

[73] T. McReynolds and D. Blythe. Advanced Graphics Programming Techniques Using OpenGL. Course presented at ACM SIGGRAPH 1999. http://www.opengl.org/developers/code/sig99.

[74] F. Metelli. The perception of transparency. *Scientific American*, 230(4):90–98, 1974.

[75] A. Mohammad-Djafari. Image reconstruction of a compact object from a few number of projections. In *Proc. International Conference on Signal and Image Processing 1996*, pages 325–329, 1996.

[76] K. Mueller, T. Möller, and R. Crawfis. Splatting without the blur. In *Proc. IEEE Visualization 1999*, pages 363–370, 1999.

[77] K. Mueller, N. Shareef, J. Huang, and R. Crawfis. IBR assisted volume rendering. In *Proc. IEEE Visualization 1999*, pages 5–8, 1999.

[78] K. Nakayama, S. Shimojo, and V. Ramachandran. Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, 19:497–513, 1990.

[79] S. Narasimhan and S. Nayar. Chromatic framework for vision in bad weather. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 2000*, pages 598–605, 2000.

[80] S. Nayar and S. Narasimhan. Vision in bad weather. In *Proc. IEEE International Conference On Computer Vision 1999*, pages 820–827, 1999.

[81] R. Neal. Probabilistic inference using Markov chain Monte Carlo methods. Technical Report CRG-TR-93-1, University of Toronto, 1993.

[82] D. Nguyen, R. Fedkiw, and H. Jensen. Physically based modeling and animation of fire. In *Proc. ACM SIGGRAPH 2002*, pages 721–728, 2002.

[83] T. Nishita, T. Sirai, K. Tadamura, and E. Nakamae. Display of the earth taking into account atmospheric scattering. In *Proc. ACM SIGGRAPH 1993*, pages 175–182, 1993.

[84] NIST Fire Dynamics Simulator (FDS) and Smokeview. http://fire.nist.gov/fds.

[85] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 1999.

[86] S. Panchev. *Random functions and turbulence*. Pergamon Press, Oxford, 1971.

[87] K. Perlin. An image synthesizer. In *Proc. ACM SIGGRAPH 1985*, pages 287–296, 1985.

[88] A. D. Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*, 14:132–137, 1995.

[89] M. Pollefeys, R. Koch, M. Vergauwen, and L. V. Gool. Automated reconstruction of 3d scenes from sequences of images. *Journal Of Photogrammetry And Remote Sensing*, 55(4):251–267, 2000.

[90] P. N. Price, M. L. Fischer, A. J. Gadgil, and R. G. Sextro. An algorithm for real-time tomography of gas concentrations, using prior information about spatial derivatives. *Atmospheric Environment*, 35(16):2827–2835, 2001.

[91] G. Ramponi and S. Carrato. An adaptive irregular sampling method for progressive transmission. In *Proc. IEEE International Conference on Image Processing 1998*, pages 747–751, 1998.

[92] F. Rashid-Farrokhi, K. J. R. Liu, C. A. Berenstein, and D. Walnut. Wavelet-based multiresolution local tomography. *IEEE Transactions on Image Processing*, 6(10):1412–1430, Oct. 1997.

[93] W. T. Reeves. Particle systems–a technique for modeling a class of fuzzy objects. In *Proc. ACM SIGGRAPH 1983*, pages 91–108, 1983.

[94] S. Roy and I. J. Cox. A maximum-flow formulation of the N-camera stereo correspondence problem. In *Proc. IEEE International Conference On Computer Vision 1998*, pages 492–502, 1998.

[95] H. Rushmeier, A. Hamins, and M. Y. Choi. Volume rendering of pool fire data. *IEEE Computer Graphics and Applications*, 15(4):62–66, July 1995.

[96] H. Rushmeier and K. Torrance. The zonal method for calculating light intensities in the presence of a participating medium. In *Proc. ACM SIGGRAPH 1987*, pages 293–302, 1987.

[97] B. Sahiner and A. E. Yagle. Image reconstruction from projections under wavelet constraints. *IEEE Transactions on Signal Processing*, 41(12):3579–3584, 1993.

[98] G. Sakas. Modeling and animating turbulent gaseous phenomena using spectral synthesis. *The Visual Computer*, 9(4):200–212, Jan. 1993.

[99] K. Sauer and C. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Transactions on Signal Processing*, 41(2):534–548, 1993.

[100] A. Sch odl, R. Szeliski, D. H. Salesin, and I. Essa. Video textures. In *Proc. ACM SIGGRAPH 2000*, pages 489–498, 2000.

[101] A. Schwarz. Multi-tomographic flame analysis with a schlieren apparatus. *Measurement Science and Technology*, 7:406–413, 1996.

[102] S. M. Seitz and C. R. Dyer. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35(2):151–173, 1999.

[103] J. Shade, S. J. Gortler, L. wei He, and R. Szeliski. Layered depth images. In *Proc. ACM SIGGRAPH 1998*, pages 231–242, 1998.

[104] A. R. Smith. Alpha and the history of digital compositing, Microsoft Technical Memo #7, August 15, 1995.

[105] A. R. Smith and J. F. Blinn. Blue screen matting. In *Proc. ACM SIGGRAPH 1996*, pages 259–268, 1996.

[106] J. Stam. Stochastic rendering of density fields. In *Proc. Graphics Interface 1994*, pages 51–58, 1994.

[107] J. Stam. *Multi-Scale Stochastic Modelling of Complex Natural Phenomena*. PhD thesis, University of Toronto, Department of Computer Science, 1995.

[108] J. Stam. Stable fluids. In *Proc. ACM SIGGRAPH 1999*, pages 121–128, 1999.

[109] J. Stam and E. Fiume. Turbulent wind fields for gaseous phenomena. In *Proc. ACM SIGGRAPH 1993*, pages 369–376, 1993.

[110] J. Stam and E. Fiume. Depicting fire and other gaseous phenomena using diffusion processes. In *Proc. ACM SIGGRAPH 1995*, pages 129–136, 1995.

[111] J. Stam and E. Languenou. Ray tracing in non-constant media. In *Proc. Eurographics Workshop on Rendering 1996*, pages 225–234, 1996.

[112] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition 2000*, pages 246–253, 2000.

[113] R. Szeliski and P. Golland. Stereo matching with transparency and matting. In *Proc. IEEE International Conference On Computer Vision 1998*, pages 517–524, 1998.

[114] J. Takahashi, H. Takahashi, and N. Chiba. Image synthesis of flickering scenes including simulated flames. *IEICE Transactions on Information and Systems*, E80-D(11):1102–1108, Nov. 1997.

[115] Y. Takai, K. Ecchu, and N. K. Takai. A cellular automaton model of particle motions and its applications. *The Visual Computer*, 11(5):240–252, 1995.

[116] R. H. Tourin. *Spectroscopic Gas Measurement*. Elsevier Publishing Company, 1961.

[117] P. van Laarhoven and E. Aarts. *Simulated Annealing: Theory and Applications*. Kluwer Academic Publishing, Dordrecht, 1987.

[118] R. Venkataramani and Y. Bresler. Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals. *IEEE Transactions on Signal Processing*, 49(10):2301–2313, Oct. 2001.

[119] S. Webb. SPECT reconstruction by simulated annealing. *Physics in Medicine and Biology*, 34(4):259–281, 1989.

[120] F. J. Weinberg. *Optics of Flames: Including methods for the study of refractive index fields in combustion and aerodynamics*. Butterworths, London, 1963.

[121] L. Westover. Footprint evaluation for volume rendering. In *Proc. ACM SIGGRAPH 1990*, pages 367–376, 1990.

[122] Y. Wexler, A. W. Fitzgibbon, and A. Zisserman. Bayesian estimation of layers from multiple images. In *Proc. European Conference on Computer Vision 2002*, volume 3, pages 487–501, 2002.

[123] G. Wolberg. *Digital Image Warping*. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[124] Y. Xu, K. L. Yung, Z. Wu, and T. Li. Computer simulation on reconstruction of 3-D flame temperature distribution. *Computational Mechanics*, 26:478–482, 2000.

[125] W. Xue, W. Donglou, and P. Gongpei. Use of Moire tomography to measure the temperature field of the flame of a pyrotechnical composition from its infrared radiation. *Combustion, Explosion, and Shock Waves*, 37(4):440–442, 2001.

[126] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. European Conference on Computer Vision 1994*, volume 2, pages 151–158, 1994.

[127] W. Zhu, Y. Wang, and J. Zhang. Total least-squares reconstruction with wavelets for optical tomography. *Journal of the Optical Society of America, A: Optics and Image Science*, 15(10):2639–2659, Oct. 1998.

[128] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin. Environment matting and compositing. In *Proc. ACM SIGGRAPH 1999*, pages 205–214, 1999.