1 Gradients of Similarity Measures

In the following we state the gradients of the similarity measures - mutual information, correlation ratio, and correlation coefficient. This completes the article and gives the reader all the necessary information for implementing an efficient gradient-based optimization of the multivariate cost function. A good reference, with detailed deduction of the derivatives, is the work of Hermosillo et al. [1].

The gradient that we show in the following was in the article denoted by

$$J_{SM,j}(x) = \frac{\partial SM(I_j(x), I_j(x \circ \exp(h))); p)}{\partial h}\bigg|_{h=0}$$

$$= \nabla SM(I_j(x), I_j(x); p).$$

We consider $I_i$ to be the fixed and $I^j_i$ to be the moving image. We add the symbol $\downarrow$ to indicate the moving image, with respect to which we derive. We define the following auxiliary variables (mean, variance) with $i_1$ an intensity in $I_i$ and $i_2$ an intensity in $I^j_i$

$$\mu_1 = \int i_1p(i_1)di_1$$

$$\mu_2 = \int i_2p(i_2)di_2$$

$$\mu_{2|1} = \int i_2p(i_1, i_2)\frac{di_1}{p(i_1)}$$

$$v_1 = \int i_1^2p(i_1)di_1 - \mu_1^2$$

$$v_2 = \int i_2^2p(i_2)di_2 - \mu_2^2$$

$$v_{1,2} = \int i_1i_2p(i_1, i_2)d(i_1, i_2) - \mu_1 \cdot \mu_2$$

$p(i_1)$ the probability for intensity $i_1$ in $I_i$ and $p(i_1, i_2)$ the joint probability.

We apply a kernel-based Parzen window method for the non-parametric PDF estimation [2], working with Gaussian kernels. For the estimation of the kernel window size, several methods were proposed. One common technique considers the maximization of a pseudo likelihood [3], which has the drawback of a trivial maximum at zero [4]. Instead, a leave-one-out strategy was proposed [3], [1]. The reported Parzen window kernel size that led to best results was five [4], which we adopted in our experiments.

1.1 Mutual Information

The formula for mutual information is

$$MI(I_i, I^j_i) = H(I_i) + H(I^j_i) - H(I_i, I^j_i)$$

$$= \int_{R^2} p(I_i, I^j_i) \log \frac{p(I_i, I^j_i)}{p(I_i) \cdot p(I^j_i)}$$

with $H$ the entropy. The derivation is

$$\nabla MI(I_i, I^j_i) = G_\Psi \frac{1}{|\Omega|} \left( \frac{\partial}{\partial I_i} p(I_i, I^j_i) - \frac{\partial}{\partial I^j_i} p(I_i, I^j_i) \right)$$

with the Gaussian $G_\Psi$ and the image grid $|\Omega|$.

1.2 Correlation Ratio

The formula for correlation ratio is

$$CR(I_i, I^j_i) = 1 - \frac{\mathbb{E}(\text{Var}(I^j_i | I_i))}{\text{Var}(I^j_i)}.$$  

The derivation is

$$\nabla CR(I_i, I^j_i) = G_\Psi \frac{\mu_2 - \mu^{2|1} + CR(I_i, I^j_i) \cdot (i_2 - \mu_2)}{\frac{1}{2} \cdot v_2 \cdot |\Omega|}.$$  

1.3 Correlation Coefficient

The formula for correlation coefficient is

$$CC(I_i, I^j_i) = \frac{\langle I_i - \mu_1 \rangle \langle I^j_i - \mu_2 \rangle}{\sqrt{v_1 \cdot v_2}}$$

and its derivation

$$\nabla CC(I_i, I^j_i) = -\frac{2}{|\Omega|} \left[ \frac{v_{1,2}}{v_2} \left( \frac{i_1 - \mu_1}{v_1} \right) + CC(I_i, I^j_i) \left( \frac{i_2 - \mu_2}{v_2} \right) \right].$$
REFERENCES


