Matching Interest Points Using Projective Invariant Concentric Circles

Han-Pang Chiu Tomás Lozano-Pérez
Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology

***Abstract* — We present a new method to perform reliable matching between different images. This method exploits a projective invariant property between concentric circles and the corresponding projected ellipses to find complete region correspondences centered on interest points. The method matches interest points allowing for a full perspective transformation and exploiting all the available luminance information in the regions. Experiments have been conducted on many different data sets to compare our approach to SIFT local descriptors. The results show the new method offers increased robustness to partial visibility, object rotation in depth, and viewpoint angle change.**

***Index Terms* — Image Matching, Interest Point, Local Descriptors.**

I. INTRODUCTION

Image matching is an essential aspect of many approaches to problems in computer vision, including object recognition [4, 6, 12], stereo matching [16, 20], motion tracking [3, 7], and segmentation [8]. A prominent approach to image matching has consisted of identifying "interest points" in the images, finding photometric descriptors of the regions surrounding these points and then matching these descriptors across images [9, 10, 11, 12, 13, 16]. The development of these methods has focused on finding interest points and local image descriptors that are invariant to changes in image formation, in particular, invariant to affine transformations of the image regions.

In this paper, we explore an alternative method for matching interest points that does not rely on local photometric descriptors but, instead, involves building a direct correspondence of regions centered on the interest points. It is widely held that photometric descriptors should be more robust to matching under imaging variation than direct image matching. This paper describes an approach to matching regions centered on interest points that allows for affine, or even projective, transformations between the images. We find that the region matching approach, even when limited to affine transformations, performs better than matching based on the popular SIFT descriptors [12]. The extension to perspective transformations produces a small gain in accuracy.

Concretely, we assume that a set of "interest points" has been identified in each image and that these image points are likely to correspond to different views of the same points in the scene. The problem is then to decide which pairs of interest points in fact correspond to the same scene point. This decision will be based on the similarity of the image regions around candidate pairs of interest points.

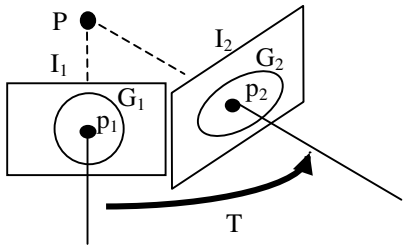


Figure 1. Characterizing regions for image matching.

This situation is illustrated in Figure 1, which shows one image point p_1 in an initial (reference) image I_1 and another image point p_2 in a second (transformed) image I_2. We want to compute a cost of the match of p_1 to p_2 based on the similarity of the colors in the image patches around them. We will assume that p_1 and p_2 correspond to point P in the scene, which is on a 3D planar surface in the scene. We want to compute a cost that is insensitive to the viewpoint change between the images. Ideally, we want the cost to be defined as follows:

$$Cost = \min_{T} \sum_{p \text{ in } G_1} (I_1(p) - I_2(T \bullet p))^2 \qquad (1)$$

where T is any perspective transformation. In our approach, we will define the region G_1 to be a set of concentric circles around p_1 and attempt to find a set of nested ellipses around p_2 that minimize the image differences and that satisfy the conditions required for a projective transformation. This set of nested ellipses defines G_2.

A. Previous Work

Interest points, local features for which the signal changes two-dimensionally, have been shown to be more robust to changing viewpoint, lighting and partial visibility than other local feature types. Local photometric descriptors computed at interest points have proved to be

This work was supported in part by the Singapore-MIT Alliance and by the Defense Advanced Research Projects Agency.

Both authors are with the M.I.T. Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA (email: chiu@mit.edu, tlp@mit.edu).

very successful in applications such as matching and recognition [9, 12, 13]. Schmid and Mohr [9] proposed one of the first matching methods based on interest points extracted by the Harris detector [10]. Dufournaud et al. [11] match images by computing interest points and descriptors at different scales but the complexity is prohibitive. To cut down the complexity, Mikolajczyk and Schmid [13] apply scale selection [14] to select characteristic points. Lowe [12] identifies stable points in scale-space based on local extrema of difference-of-Gaussian filters. Those approaches compute a cost of the match of p_1 to p_2 as shown in Fig 1 based on the similarity of the local descriptors sampled from small image patches around them. However, those approaches do not handle the case that the transformed image is taken from a substantially different 3D viewpoint because they are not fully invariant to affine and perspective transformations [15].

There has been much recent work on extending local features to be invariant to affine transformations, because a perspective transformation of a smooth surface can be locally approximated by an affine transformation. The idea is to compute local descriptors from constructed "affine invariant image regions" which define the regions G_1 and G_2 around interest points for matching as shown in Fig. 1. For example, Mikolajczyk and Schmid [15] detect affine invariant points with associated affine invariant regions. Tuytelaars and Van Gool [16] have developed affine invariant regions associated with interest points based on image intensity. The repeatability of point detection is high in these methods even when the transformed image is taken across a substantial range of viewpoints. But the performance of local descriptors computed from the invariant image region depends on the accuracy of region detection, which decreases for significant geometric transformations.

Brown and Lowe [4] developed one method that is invariant to perspective transformations, not locally approximating with an affine transformation. They use groups of interest points to form perspective invariant regions but without central interest points. Local descriptors are formed by re-sampling the image relative to canonical frames defined by the groups of interest points. But this method is quite sensitive to inaccuracy in point detection.

Image matching based on affine invariant interest point detectors and local photometric descriptors has been shown to work well in the presence of changes in viewpoint, partial visibility and extraneous features. However, not surprisingly, the accuracy of matching decreases with substantial changes in viewpoint and also when the regions around the interest points lack distinctive textures. Two possible contributing factors for this are (1) the fact that an affine approximation to the actual perspective transformation is used and (2) the sparsity of the local descriptors used in describing the regions. In this paper, we explore an approach to matching interest points that allows for a full perspective transformation and that exploits all the available luminance information in the regions.

B. *Our Approach*

Our approach aims to find an explicit correspondence for all the pixels in the regions around a pair of interest points, see Fig. 2. We match a sequence of concentric circles in one image to a sequence of embedded ellipses in the other image. The basis for matching is a projective property that is invariant between concentric circles and the corresponding ellipses under any perspective transformation. The radius of the matched circles is increased until a sudden change in the intensity difference is detected, which usually signals occlusion. We address the problem of changes in illumination by normalizing intensity in the R, G, and B channels [5].

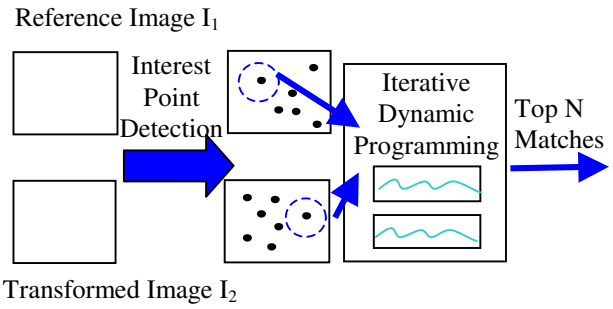


Figure 2. Overview of our method.

First we convert the images centered on each candidate pair of interest points from Cartesian coordinates (X, Y) into polar coordinate (r, θ). The problem of matching concentric circles in the reference image and the corresponding projected ellipses in the transformed image becomes that of finding the path in the polar sampled transformed image, constrained by the projective transformation equation derived in Section 2, which has the smallest color difference cost from the line in the polar sampled reference image as shown in Fig 3.

We proposed an iterative dynamic programming algorithm to find the path constrained by the projective transformation equation. It matches the regions based on repeatedly solving three-dimensional dynamic programming problems and estimating the parameters of the projective transformation equation. It can also detect the sudden change in intensity difference during the process of matching, for occlusion detection. The best N matched points with lowest average cost are returned. The details of our matching procedure are given in Section 3.

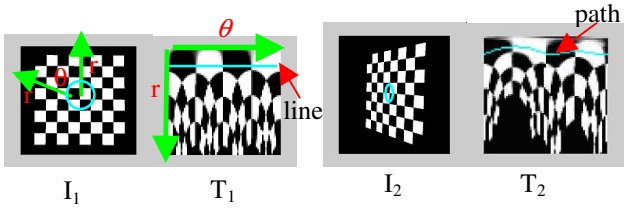


Figure 3: The reference image I_1, the transformed image I_2, and the polar sampled images T_1 and T_2 generated from them respectively. Pixels in T_1 and T_2 are indexed by radius r and angle θ. The circle in I_1 and the corresponding projected ellipse in I_2 become the line in T_1 and the path in T_2 respectively.

Experiments have been conducted to compare this approach (and an affine version) to matching using the SIFT local descriptors [12], which perform best among existing local descriptors [2]. To make the comparison fair, all three methods use the same input set of "interest points" obtained by one recent affine invariant detector proposed by Mikolajczyk and Schmid [15]. This method estimates the second moment matrix that can be used to normalize a region in an affine invariant way around an interest point. SIFT descriptors are computed from these affine-invariant regions around interest points and compared using Euclidean distance. The results show our approach is better at handling partial visibility and greater object rotation in depth (Section 4).

II. Ellipses and Concentric Circles

Concentric circles and the corresponding projected ellipses have been actively used for pose estimation as well as for camera calibration [17, 18]. Although ellipse-based techniques have some intrinsic properties that have been exploited in vision applications such as factory automation and automatic assembly, they have been used relatively little in general image matching.

Kim et al. [19] found an important projective invariant property of projected circle centers between reference image I_1 and transformed image I_2. Note that, under a general projective transformation, the projections of concentric circles will be ellipses with different centers. One can show that the projected concentric circle center is on the line defined by the centers of the nested ellipses resulting from the projection, see Fig. 4. The centers of resulting ellipses all map to the projected concentric circle center only under pure affine transformations.

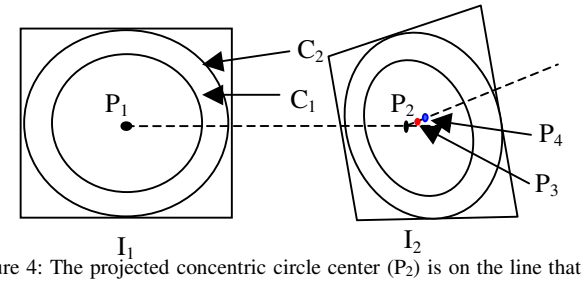


Figure 4: The projected concentric circle center (P_2) is on the line that is defined by the two ellipse centers. The ellipse with the red point center P_3 is the result of the projection of circle C_1. The ellipse with the blue point center P_4 is the result of the projection of circle C_2.

This property is consistent with the difference between the degrees of freedom in affine and projective transformation. An affine transformation has six degrees of freedom; projective transformations have eight degrees of freedom. The additional two degrees of freedom can be written as a two-component co-vector v, which in addition to the 2 by 2 affine matrix A and the offset two-component vector t can be used to define a projective transformation [1]. x_1 and x_2 are image coordinates of matched points on the circle in I_1 and the projected ellipse in I_2 respectively. It can easily be shown that v determines the translation of the projected ellipse center in I_2. If v is a zero vector, this form represents an affine transformation and the transformed circles are co-incident.

$$x_2 = \frac{Ax_1 + t}{1 + vx_1}$$

We can use this projective transformation equation to constrain the matched regions between two images. Since a circle in reference image I_1 should be an ellipse in transformed image I_2 under any projective transformation, we define that the origins of the polar coordinates in I_1 and I_2 are P_1 and P_2 respectively, thus the two translation degrees of freedom (represented by vector t) can be fixed. The point x_1 on the circle with radius R and angle θ_1 in I_1 has the form ($R\cos\theta_1$, $R\sin\theta_1$) and each point in the polar coordinate system of I_2 has the form ($r\cos\theta_2$, $r\sin\theta_2$). So the final equation to constrain the projection of concentric circles in I_2 becomes.

$$[r\cdot\cos\theta_2 \quad r\cdot\sin\theta_2]^T = \frac{A\cdot[R\cdot\cos\theta_1 \quad R\cdot\sin\theta_1]^T}{1 + v\cdot[R\cdot\cos\theta_1 \quad R\cdot\sin\theta_1]^T}$$

$$A \in \Re^{2\times 2}, v \in \Re^2 \qquad (2)$$

There are ten parameters in this equation (2). But if P_1 and P_2 are obtained by an affine invariant point detector, the detector already estimated the second moment matrices A_1 and A_2. So, we can normalize elliptical regions E_1 and E_2 in an affine invariant way around center points P_1 and P_2 respectively. Then we can set the affine matrix $A = A_2A_1^{-1}$ which projects elliptical region E_1 onto E_2 as shown in Figure 5. Since A_1 and A_2 are estimated from small affine-invariant regions, which are not influenced much by the non-affine vector v ($v\cdot[R\cdot\cos\theta_1 \quad R\cdot\sin\theta_1]^T$ is nearly zero when R is very small), this is likely to be a good guess for A. So in our implementation, we just estimate the two non-affine parameters in v based on A and a set of known (R, r, θ_1, θ_2), that is, the set of points (r, θ_2) on an ellipse in the transformed image corresponding to the set of points on a circle of radius R and associated angle θ_1 in the original image.

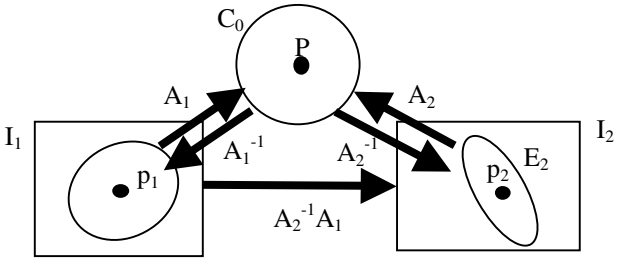


Figure 5: Both points p_1 and p_2 correspond to P, the center of 3D circle C_0. A_1 and A_2 are the second moment matrices to normalize elliptical regions E_1 and E_2 to C_0 in an affine invariant way around center points P_1 and P_2. The affine transformation matrix A from E_1 to E_2 will be the product of A_2^{-1} and A_1.

III. Matching and Recognition Method

For any two images we want to match, we use the same local invariant interest point detector [15] proposed by Mikolajczyk and Schmid to obtain the set of points in both images. Then for each pair of interest points that we want to match, we generate polar-sampled (m angles) templates T_1 and T_2 originated from the interest points in the reference image I_1 and transformed image I_2 respectively as shown in Fig. 6.

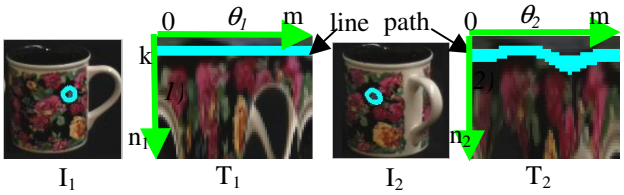


Figure 6: The reference image I_1, the transformed image I_2, and the polar sampled templates T_1 and T_2 generated from them respectively. The narrow horizontal line in T_1 and the corresponding circle in I_1 are marked. The resulting path generated by dynamic programming method in T_2 and the corresponding projected ellipse in I_2 are marked. There are n_1 rows and n_2 rows in T_1 and T_2 respectively.

Note that as the figure below makes clear, corresponding columns in the template cannot be compared directly. There is an unknown stretching along the θ axis as well as an unknown rotation.

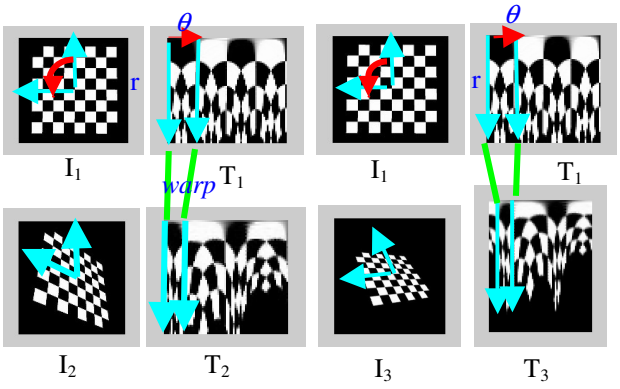


Figure 7: Two corresponded rays in reference image I_1 and transformed image I_2 are marked respectively. The warping situation exists in the polar sampled templates T_1 and T_2 generated from I_1 and I_2. Polar sampled templates T_1 and T_3 generated from I_1 and I_3 can't be compared directly due to an unknown image rotation.

To address these issues, for each pair of matched interest points, we run a dynamic programming algorithm to match polar sampled templates T_1 and T_2 generated from reference image I_1 and transformed image I_2 respectively. We take one narrow horizontal line at row k in T_1 (corresponding to a circle of radius R in I_1) and find the corresponding path in T_2 (corresponding to an ellipse in I_2) by the three-dimensional warping algorithm formulated below. Note that the position row ρ of a path can move up or down at each step and that strip column C_1 on T_1 could warp to more than one strip columns on T_2, and vice versa.

$$cost(m,m) = \min_{1\le\rho\le n_2}[cost(m,m,\rho)]$$

$$cost(C_1,C_2,\rho) = \min(a_1,a_2,a_3) + d(C_1,C_2,\rho)$$

$$cost(1,1,1) = d(1,1,1)\ \ 1\le C_1,C_2\le m\ \ 1\le\rho\le n_2$$

$$a_1 = \min[cost(C_1-1,C_2,\rho-1), cost(C_1-1,C_2,\rho), cost(C_1-1,C_2,\rho+1)]$$

$$a_2 = \min[cost(C_1,C_2-1,\rho-1), cost(C_1,C_2-1,\rho), cost(C_1,C_2-1,\rho+1)]$$

$$a_3 = \min[cost(C_1-1,C_2-1,\rho-1), cost(C_1-1,C_2-1,\rho), cost(C_1-1,C_2-1,\rho+1)]$$

$$d(C_1,C_2,\rho) = (T_1(k,C_1) - T_2(\rho,C_2))^2$$

This is started with one particular column on T_1 and different starting columns on T_2 so as to handle image rotation. The one with lowest cost(m,m) is the starting column of T_2 we want. Ideally, we would want to constrain the dynamic programming process to produce only valid ellipses but we do not have an efficient algorithm for this. Instead, after the path is found in the polar sampled template T_2, we obtain a set of (r, θ_2) from the matched pixels along the path, as shown in Fig. 6, and can estimate the non-affine parameters v by minimizing the error function derived from the constraint equation (2) and the measured (R, r, θ_1, θ_2, A).

Since the transformation parameters of adjacent projected ellipses are identical, we take another horizontal line at row K where K > k in T_1 and predict the mapped path (r, θ_2 **:** the row r along each column θ_2) on T_2 with the estimated projective transformation constraint equation (2) since parameters of v and the warping result on T_1 and T_2 are known. Then we can apply the dynamic warping algorithm again, with this circle of larger radius. But $d(C_1,C_2,\rho)$ is now based both on the color difference and on the error between the matched column position and the predicted column position as follows:

$$d(C_1,C_2,\rho) = \sqrt{(T_1(k,C_1) - T_2(\rho,C_2))^2} + \sqrt{(\mathrm{Predict}(C_2) - \rho)^2}$$

where Predict(C_2) is the predicted path row r along column C_2 on T_2. This process is repeated iteratively with K, the input horizontal line in T_1, increasing. The two parameters of vector v in the constraint equation (2) are refined during this iterative process. Thus the matched region is grown by iteratively increasing the radius of the original circle.

If (I_1 (θ_1, R)-I_2 (θ_2, r)) of any angle θ_1 along the trajectory is bigger than some pre-specified threshold, we mark the position with this particular angle θ_2 along the trajectory of the ellipse as being occluded as shown in Fig. 8. We keep track of what fraction of the rays radiating from the interest point is marked occluded and stop the iterative process if this fraction exceeds 1/3. This process does a good job of stopping the growing of the matched regions in the presence of partial occlusion.

The top N candidate pairs of matched interest points with the lowest average cost will be preserved. They are the pairs of matched points p_1 and p_2 that we will return. The matched regions around p_1 and p_2 in the reference image I_1 and the transformed image I_2 respectively after the process is the final result of image matching in our method.

This method handles perspective transformations, exploiting the projective invariant property between concentric circles and the corresponding projected ellipses to find complete region correspondences. It matches images well even in the presence of large rotations in depth. The iterative process can match larger image regions around the interest points and is good at dealing with occlusion.

I_1 I_2

Figure 8: The concentric circular region in the reference image I_1 and the elliptical region in the transformed image I_2 that is the result of the projection of the concentric circles. The lighter colored points indicate possible occlusion positions we detected along the trajectory of the projected ellipse.

IV. Experiment Results

In this section we report the results of experiments with the matching method described in Section 3. The results of this method are compared with those from the SIFT local descriptors [12][1]. We also compare the results with a variant of our method that limits itself to affine transformations, which we will refer to as affine invariant concentric circles.

To make the comparison fair, all three methods use the same input set of "interest points" obtained by the affine invariant points method of Mikolajczyk and Schmid [15]. This detector estimates the second moment matrix, which can be used to normalize regions in an affine invariant way around interest points. SIFT descriptors are computed from these affine-invariant regions and compared with the Euclidean distance for matching. We have chosen to base our comparison on SIFT descriptors since they performed best among existing local photometric descriptors in Mikolajczyk and Schmid's evaluation [2].

The affine approximation version of our method is used to assess the impact of the various components of our method. It doesn't need to estimate the non-affine co-vector v; it just uses the affine matrix A derived from the affine invariant detector as shown in Fig. 5. So the region matching process around points becomes a simple iterative one-dimensional dynamic warping algorithm to handle the image stretching issue as shown in Fig. 7. Comparing the results of the two versions we can assess the benefit from projective rather than affine invariant concentric circles. It also helps us discriminate the impact of detailed pixel matching versus using the SIFT local descriptors since it uses the same affine transformation obtained by Mikolajczyk and Schmid's method.

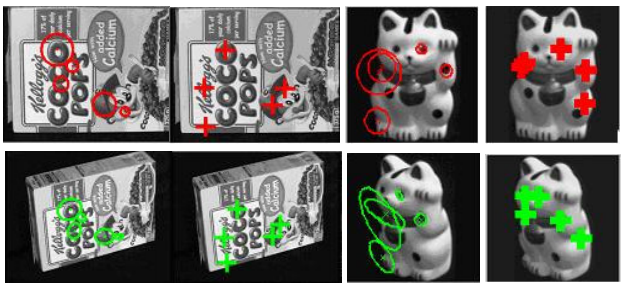
[1] The code for the point detector and local SIFT descriptors were obtained from www.inrialpes.fr/lear/people/Dorko/downloads.html.

First, we conducted experiments on five different data sets of images to address several issues that can affect the performance of image matching methods, including significant geometric transformation, texture and occlusion. Each data set contains 10 image pairs collected from the Internet. All three methods return the best N matches for each image pair (N is usually 5 or 10, depending on the size of the image). We report the accuracy rate (percentage of correct matches among the best N matches) to evaluate the performance of each method. The performance of each method on the data sets is displayed in Table 1. Our initial implementation takes about 10.7 seconds to process a pair of 800*640 images on a standard 2.52 GHz Pentium PC.

A. Data Set Type	**Projective Invariant Concentric Circles**	**Affine Invariant Concentric Circles**	**SIFT**
Small rotation	0.94/50	0.92/50	0.90/50
Large rotation	0.58/50	0.50/50	0.36/50
Distinctive texture	0.91/100	0.91/100	0.81/100
Similar texture	0.72/100	0.70/100	0.68/100
Occlusion	0.82/50	0.80/50	0.72/50

Table 1: Accuracy rate/total number of returned matches of the three methods on the five data sets.

All methods work well on the dataset where the object in the image is rotated about 30 degrees in depth, Fig. 9(a)(b). To evaluate the performance for significant geometric transformations, we used images where objects are rotated from 40 to 60 degrees in depth, Fig 9(c)(d). Note that all the methods work less well for these challenging matching tasks, when the rigid transformation is large.



(a) (b) (c) (d)

Fig. 9. (a) All five matches returned by our method are correct. (b) All five matches returned by SIFT descriptors are correct. (c) All five matches returned by our method are correct. (e) Only three of five matches returned by SIFT descriptors are correct.

In images with distinctive texture patterns, SIFT local descriptors occasionally accept some incorrect matches due to occlusion or to large change in viewpoint. Matching with projective or affine invariant concentric circles is more robust, as illustrated in Fig. 10(a)(b). Presumably, this is

because we incorporate more photometric information around matched points to verify the matches. For images with less-distinctive texture, the SIFT descriptors have a high error rate, presumably because the descriptors are not sufficiently distinctive [15]. The results of our methods are a bit better but it is still susceptible to errors in detection of interest points. If the repeatability of point detection is high both the methods will be nearly perfect, Fig. 10(c)(d). But if not, Fig 11, none of the methods can get good results. This remains a fundamental limitation of interest point matching.

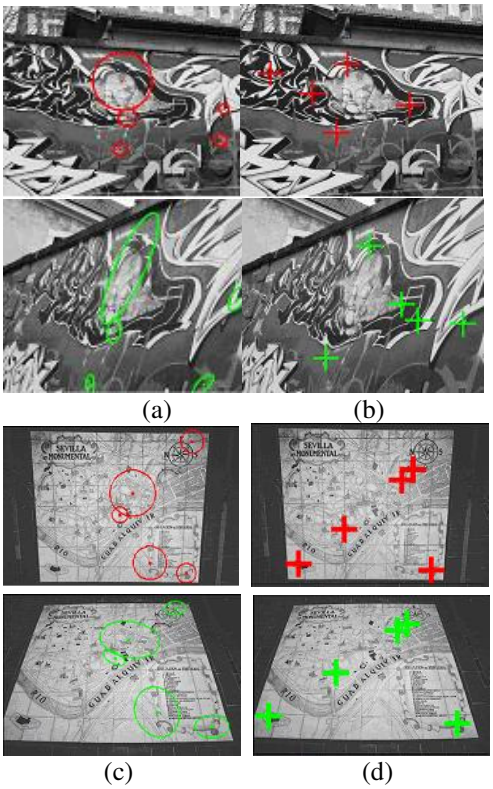


Fig. 10. (a) All five matches returned by our method are correct. (b) Two of five matches returned by SIFT descriptors are the matches that the correct answer does not exist. (c) All five matches returned by our method are correct. (d) All five matches returned by SIFT descriptors are correct.

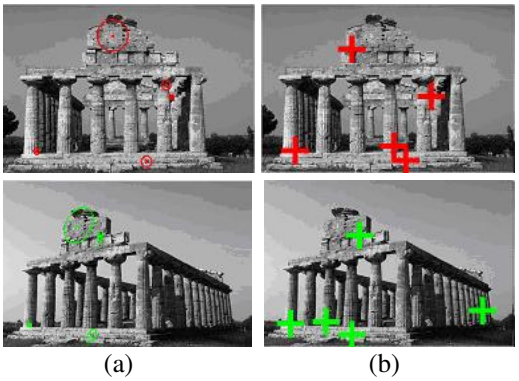


Fig. 11. (a) Only the first three of five matches returned by our method matches are correct. (b) Three of five matches returned by SIFT descriptors are wrong.

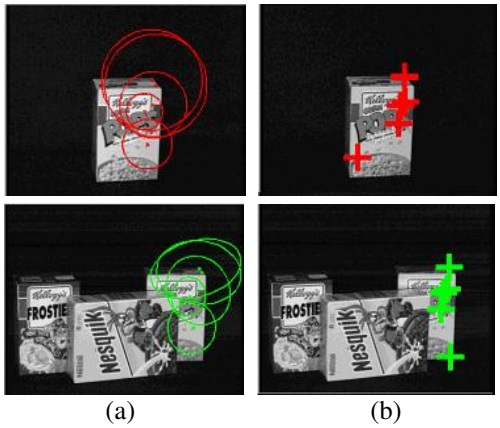


Fig. 12. (a) All five matches returned by our method are correct. (b) One of four matches returned by SIFT descriptors is wrong due to occlusion.

Occlusion is an important issue in image matching tasks. To assess the impact on our method, we used some pairs of images where the object in the reference image is partially covered by other objects in the transformed image. The results of our methods in these cases are better than when using SIFT descriptors, as illustrated in Fig. 12. SIFT descriptors generate mismatches if the texture around an incorrectly matched point in the transformed image is similar to that of the correct match, which is occluded. It rarely happens in our methods because the matched region is extended to verify the match and we detect the sudden change in intensity difference due to the occlusion.

Although our methods perform better than SIFT descriptors in all of these data sets, the gain from matching with projective invariant concentric circles versus affine ones is slight. One reason is that the correct matching points are often not included in the input set of interest points. When the viewpoint change is large or the textures not distinctive, the repeatability rate of point detection is not high. This is not surprising since the detector obtains interest points in an affine-invariant way, not projective invariant, so many initial points do not converge during the affine normalization process.

We also compared the three methods on the data sets in Mikolajczyk and Schmid's evaluation of local descriptors [2]. These data sets involve highly textured images where point detection repeatability is high. There are four data sets to address image rotation, scale changes, affine transformation and illumination changes. Each data set contains 3 image pairs. We evaluate performance of each method by computing accuracy rate; note that N varies from 10 to 60. The performances of the three methods are similar except on the affine transformation data set where the viewpoint of the camera is changed by substantial amounts as shown in Fig. 13. Once again matching with projective invariant concentric circles performs best in this particular data set. Affine invariant concentric circles come a close second and performs better than SIFT descriptors.

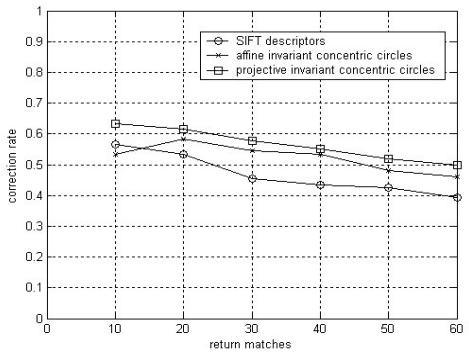


Fig. 13. The performance measure (accuracy rate versus top N returned matches) of three methods on the data set involving viewpoint changes.

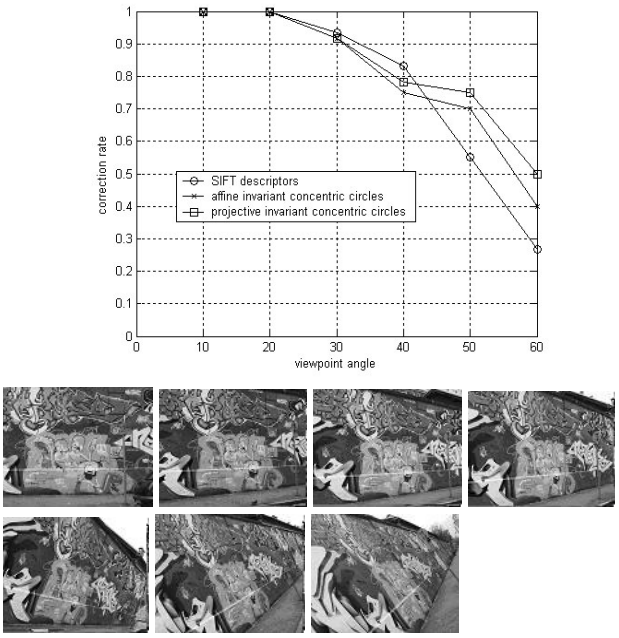


Fig. 14. The performance measure (accuracy rate versus viewpoint angle) of three methods over one image sequence with known viewpoint angle.

To demonstrate the influence of the viewpoint changes between matching with a projective transformation and an affine transformation, we experiment with several image sequences with known viewpoint angle. The accuracy rate is the percentage of correct matches among the top 60 returned matches. We found when the viewpoint angle is less than 40 degrees, all three methods work well, sometimes SIFT descriptors are slightly better. After 40 degrees, the performance of all the methods drops substantially, but projective invariant concentric circles outperforms the other two methods. One representative example is shown in Fig. 14.

V. CONCLUSIONS

In this paper we present a new method to perform reliable matching between different images. Our method constructs detailed pixel level matches between regions rather than relying on local photometric descriptors. We also showed how to exploit an intrinsic property of projected concentric circles via a constraint projective transformation equation. This enables the method to deal with perspective transformations, not only affine transformations. It is more robust than previous methods to a variety of textures and to occlusion because it incorporates more luminance information around the interest points and because it finds a more detailed region correspondence. Experiments have been conducted on many different data sets to compare our approach to SIFT descriptors on affine invariant regions around interest points [15]. The results showed the new method offers increased robustness to partial visibility, greater object rotation in depth, and more viewpoint angle change.

Future work will look at methods for finding legal ellipses that minimize the intensity difference cost directly rather than through the two-step process (dynamic programming followed by fitting) that we use now. We also plan to work on methods that can detect interest points in a projective invariant way.

REFERENCES

[1] J. G. Semple and G. T. Kneebone, Algebraic projective geometry, Oxford University Press, 1952.
[2] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors", CVPR, 2003.
[3] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in video", ICCV, 2003.
[4] M. Brown and D. Lowe, "Invariant features from interest point groups", BMVC, 2002.
[5] B. Funt, K. Barnard, and L. Martin, "Is machine colour constancy good enough?", ECCV, 1998.
[6] B. Schiele and J. L. Crowley, "Object recognition using multidimensional receptive field histograms", ECCV, 610-619, 1996.
[7] M. La Cascia, S. Sclaroff, and V. Athitsos, "Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models", IEEE Trans. on PAMI, 22(4), 2000.
[8] R. J. Qian and T. S. Huang, "Robust and accurate image segmentation using deformable templates in scale space", ISCV, 1995.
[9] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval", IEEE Trans. on PAMI, 19(5): 530-534, 1997.
[10] C. Harris and M. Stephens, "A combined corner and edge detector", Alvey Vision Conf., 147-151, 1988.
[11] Y. Dufournaud, C. Schmid, and R. Horaud, "Matching images with different resolutions", CVPR, 612-618, 2000.
[12] D. Lowe, "Object recognition from local scale-invariant features", ICCV, 1150-1157, 1999.
[13] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points", ICCV, 523-531, 2001.
[14] T. Lindeberg, "Feature detection with automatic scale selection", IJCV, 30(2):79-116, 1998.
[15] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector", ECCV, 128-142, 2002.
[16] T. Tuytelaars and L. Van Gool, "Wide baseline stereo matching based on local, affinely invariant regions", BMVC, 412-425, 2000.
[17] Z. Chen and J. Huang, "A vision-based method for the circle pose determination with a direct geometric interpretation", IEEE Trans. on Robotics and Automation, 15(6): 1135-1140, 1999.
[18] R. Hu and Q. Ji, "Camera self-calibration from ellipse correspondences", ICRA, 2191-2196, 1999.
[19] J-S Kim, H-W Kim, and In So Kweon, "A camera calibration method using concentric circles for vision applications", The 5th Asian Conference on Computer Vision, 2002.
[20] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision", Proc. IJCAI, 674-679, 1981.