

# CONTRAST-INVARIANT FEATURE POINT CORRESPONDENCE

Ping Li<sup>1</sup>, Dirk Farin<sup>1</sup>, Rene Klein Gunnewiek<sup>2</sup> and Peter H. N. de With<sup>1,3</sup>

<sup>1</sup> Eindhoven University of Technology, <sup>2</sup> Philips Research Eindhoven, <sup>3</sup> LogicaCMG Netherlands B.V.  
Emails: {p.li, d.s.farin}@tue.nl, rene.klein.gunnewiek@philips.com, P.H.N.de.With@tue.nl

## ABSTRACT

Most existing feature-matching methods utilize texture correlation for feature matching, which is usually sensitive to contrast changes. This paper proposes a new feature-point matching algorithm that does not rely on the image texture. Instead, only the smoothness assumption, which states that the displacement field in a neighborhood is coherent (smooth), is used. In the proposed method, the collected correspondences of a group of feature points within a neighborhood are efficiently determined such that the coherence measure of the displacement field in the neighborhood is maximized. The experimental results show that the proposed method is invariant to contrast changes and significantly outperforms the conventional block-matching technique.

**Index Terms**— Stereo, correspondence, feature point, contrast, feature matching

## 1. INTRODUCTION

Finding feature-point correspondences for two images can be divided into two steps: 1) detecting the feature points in the individual images; 2) establishing correspondences between the detected feature points. This paper focuses on the second step. We assume that the feature points have been detected in two images using the well-known Harris corner detector [1].

Establishing feature correspondences between two related images, such as the members of a stereo pair or successive frames in a motion sequence, has been and continues to be a central problem in computer vision. Many feature correspondence methods have been proposed in the last two decades. They can be coarsely classified into two groups: area correlation methods [2] and cooperative methods [3]. The *area correlation methods* match the feature points by measuring the similarities of two image patches around two feature points. The image measurements such as the intensity, color, phase, etc., are aggregated over the window. Two feature points are matched if the measurements show high similarity. The *cooperative methods* perform the feature matching by cooperatively considering the disparities from neighboring feature points. Typically, both the similarity assumption of the image texture and the smoothness assumption of the displacement field are utilized. Similar to the area correlation methods, the similarity is usually measured over a window around the feature point using metrics such as correlation, sum of the squared differences (SSD), etc.

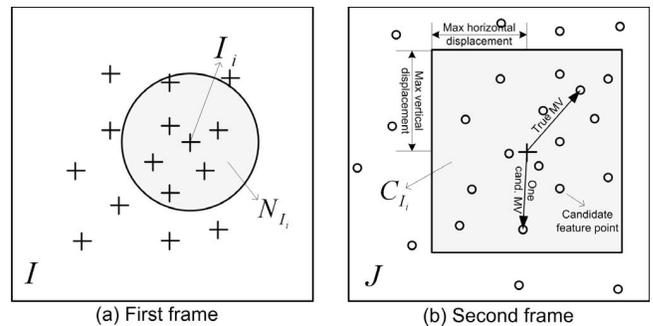
Most aforementioned methods utilize conceptually the image correlation for feature matching, which typically requires that the contrasts of the two images are constant. However, a constant contrast is difficult to maintain in practice. Even if we assume that the camera hardware is identical, for slightly different points of view, the amount of light entering the two cameras can be different, causing dynamically adjusted internal parameters such as aperture, exposure

and gain to be modified [4]. In that case, many existing methods that utilize the similarity assumption cannot work well. In [5], a feature matching method that utilizes the proximity assumption is proposed, which is invariant to contrast.

This paper proposes a new feature-matching method that works well even if the contrast changes substantially across images. Therefore we refer to it as a Contrast-Invariant Feature Matching (CIFM) algorithm. In the proposed method, only the geometry constraint stating that the displacement field in a small neighborhood is coherent<sup>1</sup> (smooth), is utilized. The collected correspondences for a group of feature points within a neighborhood are efficiently determined such that the coherence measure of the displacement field is maximized. The proposed method is thus invariant to contrast since no texture information is used for the feature matching.

The remainder of the paper is organized as follows. Section 2 introduces the problem and the notations used in this paper. Section 3 describes the proposed algorithm. Section 4 presents the experimental results and Section 5 concludes this paper.

## 2. NOTATIONS AND PROBLEM FORMULATION



**Fig. 1.** The set of feature points in the neighborhood  $N_{I_i}$  in the first frame and the set of candidate corresponding feature points  $C_{I_i}$  in the second frame for feature point  $I_i$ .

Let  $I = \{I_1, I_2, \dots, I_M\}$  and  $J = \{J_1, J_2, \dots, J_N\}$  be two sets of feature points in two related images, containing  $M$  and  $N$  feature points, respectively. Let  $(x_{I_i}, y_{I_i})$  be the coordinates of feature point  $I_i$ . For feature point  $I_i$ , we want to find its corresponding feature point  $J_j$  from all its candidates. We define the set of candidates for  $I_i$  as  $C_{I_i}$ . As shown in Fig. 1(b),  $C_{I_i}$  is defined as the set of feature points within a co-located rectangle in the second frame.

<sup>1</sup>Coherent means all points on each object surface move in the same way.

The dimensions of the rectangle define the maximum displacements allowed for the feature points.

The set of feature points within the neighborhood of  $I_i$  is denoted as set  $N_{I_i}$ . As shown in Fig. 1(a), the neighborhood of  $I_i$  is defined as a circular area with its center at  $(x_{I_i}, y_{I_i})$ . The displacement between  $I_i$  and  $J_j$  is represented by the displacement vector of  $I_i$ , i.e.,  $\vec{v}_{I_i} = (x_{\vec{v}_{I_i}}, y_{\vec{v}_{I_i}})$ . Obviously, the set of candidate feature points  $C_{I_i}$  for  $I_i$  gives rise to the set of candidate displacement vectors  $V_{I_i}$  for  $I_i$ . Thus, determining the correspondence for  $I_i$  is equivalent to selecting the corresponding feature point from  $C_{I_i}$  or selecting the corresponding displacement vector from  $V_{I_i}$ . *Point correspondences*, *displacement vectors*, and *point matches* have the same meaning throughout this paper.

### 3. ALGORITHM

The proposed algorithm avoids the use of the similarity assumption. Instead, only the geometry constraint, which is invariant to contrast and also robust to the change of camera parameters, is utilized.

#### 3.1. Coherence metric

To test if two displacement vectors  $\vec{v}_i$  and  $\vec{v}_j$  are coherent, the distance between them is efficiently computed by:

$$d_{ij} = |\vec{v}_i - \vec{v}_j| = \sqrt{(x_{v_i} - x_{v_j})^2 + (y_{v_i} - y_{v_j})^2}. \quad (1)$$

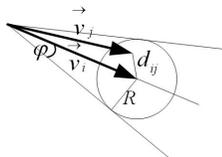
Vectors  $\vec{v}_i$  and  $\vec{v}_j$  are considered coherent if  $d_{ij}$  is smaller than a given threshold. The geometric meaning of Eq. (1) is illustrated in Fig. 2. This criteria requires that two displacement vectors in a neighborhood must align in similar directions and have similar magnitudes to be coherent.

The coherence of the displacement field within a neighborhood  $N_{I_i}$  is measured as the ratio between the number of the coherent displacement vectors found in  $N_{I_i}$  and the number of the feature points in  $N_{I_i}$ . This ratio is denoted by  $C(N_{I_i})$  and can be computed as:

$$C(N_{I_i}) = \frac{\sum_{I_k \in N_{I_i}} f_{I_k}}{n(N_{I_i})}, \quad (2)$$

where  $n(N_{I_i})$  is the number of feature points in  $N_{I_i}$ ;  $f_{I_k}$  is a binary indicator variable, indicating whether the displacement vector of feature point  $I_k$  is coherent with the reference displacement vector.

As stated by the smoothness assumption, the displacement field within a small neighborhood should be continuous. This means  $C(N_{I_i})$  should be as high as possible. Thus, the problem to determine the correspondences for each feature point  $I_k \in N_{I_i}$  is converted into selecting a displacement vector  $\vec{v}_{I_k} \in V_{I_k}$  for  $I_k$  such that  $C(N_{I_i})$  is maximized.



**Fig. 2.** Two displacement vectors  $\vec{v}_i$  and  $\vec{v}_j$  are considered coherent if their distance  $d_{ij}$  is smaller than a given threshold  $R$ .

The proposed coherence metric is similar to the rigid-motion model while allowing a certain degree of deviation of the displacement vectors. However, note our algorithm does not enforce the smoothness of the displacement field. The smoothness assumption is only used for detecting whether a coherent displacement vector exists for a feature point. For example, for the neighborhoods crossing object boundaries, finding coherent displacement vectors satisfying the proposed coherence metric for all feature points may be difficult. In that case, we simply assume no correspondence exists for those feature points for which no coherent displacement vectors can be found. In this way, the motion discontinuity is preserved though some correspondences are lost.

#### 3.2. Computing correspondences for feature points within a neighborhood

The steps to determine the correspondences for feature points in neighborhood  $N_{I_i}$  are summarized as follows: 1) for every  $\vec{v}_{I_i} \in V_{I_i}$ , find the closest  $\vec{v}_{I_k}$  from  $V_{I_k}$  for every  $I_k \in N_{I_i}$  so that  $d_{ik}$  by Eq. (1) is smallest; 2) if  $d_{ik}$  is smaller than a threshold, set the indicator variable  $f_{I_k}$  to unity. Otherwise, set it to zero; 3) compute the coherence of the displacement field using Eq. (2); 4) the set of displacement vectors with the largest  $C(N_{I_i})$  are considered as the true correspondences if  $C(N_{I_i})$  is larger than a given threshold.

#### 3.3. Rationale of the algorithm

The algorithm described above tries to find displacement vector  $\vec{v}_{I_i}$  from  $V_{I_i}$  such that a maximum number of coherent displacement vectors can be found in a neighborhood. We now briefly explain why this maximum coherence gives the true correspondences with a high probability.

Suppose the repetition ratio of the feature points in  $N_{I_i}$  is  $\alpha(N_{I_i})$ . Also suppose we successfully find the true correspondences for all the repeated feature points. Then we can say that  $(n(N_{I_i}) \times \alpha(N_{I_i}))$  number of coherent displacement vectors are found in the neighborhood<sup>2</sup>. This means that  $C_{N_{I_i}}$  equals to  $\alpha(N_{I_i})$  in the direction of the true motion  $\vec{v}_{I_i}$ . Thus, we can assume  $(C_{N_{I_i}} = \alpha(N_{I_i}))$  when the correct motion is found.

Due to the random pattern of the texture, in other directions than the true motion, feature points usually appear randomly. The chance to find another set of coherent displacement vectors in other directions that give higher coherence is thus low. The highest coherence can be found, in most cases, only in the direction of true motion.

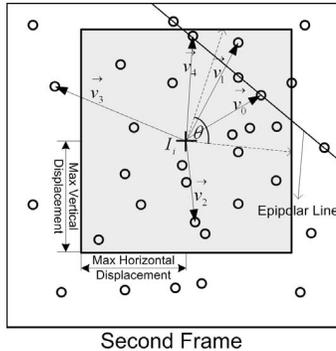
As a summary, the repetition ratio of the feature points guarantees that a coherence equal to  $\alpha(N_{I_i})$  can be found in the direction of true motion for most cases. The randomness of feature points ensures that in directions other than the true motion, the coherence is most likely lower.

Above observation holds only when the repetition ratio is not too low, so that the coherence in the direction of true motion dominates over other directions. For image areas like grass, tree, etc., where feature points rarely repeat, the probability increases that the highest coherence does not lie in the direction of the true motion. Consequently, the feature correspondences are difficult to detect for these areas.

<sup>2</sup>According to the geometry constraint, the  $n(N_{I_i}) \times \alpha(N_{I_i})$  number of true displacement vectors should be coherent. In most cases except for the neighborhoods crossing the object boundaries and the neighborhoods with significant depth changes, these coherent displacement vectors can be detected by our coherence criteria introduced in Subsection 3.1.

### 3.4. Handling low-repetition-ratio image area

To obtain robust results for the low-repetition-ratio image areas, more constraints like the epipolar constraint<sup>3</sup>, the motion constraint<sup>4</sup>, etc., can be used to reduce the search space for the true correspondence (to remove the ambiguity among several directions), which is illustrated by Fig. 3. In the figure,  $\theta$  denotes the cone in which the displacement vector of  $I_i$  should be situated. This cone can be predicted from neighboring feature points whose displacement vectors have been found. Vector  $\vec{v}_0$  is the true displacement vector of  $I_i$ ; vectors  $\vec{v}_1$ ,  $\vec{v}_2$ ,  $\vec{v}_3$  and  $\vec{v}_4$  are the four other candidate displacement vectors that give higher coherence than  $\vec{v}_0$  (due to the low repetition ratio in the area).



**Fig. 3.** Reduce the search space for feature correspondences using multiple constraints.

As shown in Fig. 3, by limiting the maximum displacement of the feature point,  $\vec{v}_3$  is eliminated; by incorporating the epipolar constraint,  $\vec{v}_1$  and  $\vec{v}_2$  are eliminated; by utilizing the motion constraint,  $\vec{v}_4$  is eliminated. Thus, the ambiguity among the several promising displacement vectors is gradually removed and the true correspondence  $\vec{v}_0$  is found. Note that ambiguities exist mostly only in the low-repetition-ratio areas. For the image areas where most feature points repeat,  $\vec{v}_0$  usually wins over all other directions and the true correspondence can be reliably detected.

In our method, only the maximum-displacement constraint is used. This first improves the robustness of our algorithm. It also reduces the size of  $C_{I_k}(I_k \in N_{I_i})$  and thus the complexity of our algorithm. The epipolar constraint is not yet utilized because it may not be applicable for some degenerate cases such as rotation-only camera motion, planar scene geometry, etc.

### 3.5. Size of the neighborhood

In the proposed algorithm, the neighborhood is chosen as the circular area around a feature point. The size of a neighborhood is determined based on the local density of the feature points. Our approach is to adjust the neighborhood size so that a fixed number of feature points are found within every neighborhood. This approach has following advantages: (1) for image areas which contain little texture and thus few feature points, a large neighborhood will be automatically selected; this is helpful for a robust feature matching

<sup>3</sup>The corresponding feature point must lie on the epipolar line.

<sup>4</sup>Motion vectors of neighboring feature points usually lie in the same direction.

in texture-scarce image areas; (2) a small neighborhood will be selected for texture-rich areas with dense feature points, which helps the accuracy of the detected feature correspondences; (3) adjusting the neighborhood size to ensure a fixed number of feature points is computationally simple.

## 4. EXPERIMENTAL RESULTS

The proposed CIFM technique has been applied to many image pairs. However, only the results for two image pairs are presented in this section since all experiments show similar results. The first pair shows a small contrast change and the second one shows a large contrast change. To evaluate the quality of the detected feature correspondences, the detected correspondences are used to compute the fundamental matrix using RANSAC [6]. All correspondences with their symmetric transfer errors smaller than one pixel are considered as true correspondences. The percentage of the true correspondences that conform to epipolar geometry over the total number of the detected correspondences is then computed. The results are then compared with those by the conventional Block Matching (BM) method. The maximum displacement is limited to 50 pixels both horizontally and vertically in both methods.

Our first experiment is on an image pair from the *castle* sequence [7] showing a small contrast change. The first row of Fig. 4 shows the correspondences obtained using the BM. By comparing Fig. 4(a) with Fig. 4(b), we see many spurious correspondences are detected by the BM.

Table 1 shows the results obtained by the BM and the CIFM on image pair 1 (IP1) and image pair 2 (IP2). In the table, *OutOfDetcd* means the percentage of the feature points that are found conforming to the epipolar constraint over the number of detected correspondences; *OutOfTotal* means the percentage of the feature points whose true correspondences have been found over the total number of feature points. As we see from the table, for the BM-IP1, among the 1332 correspondences detected out of 3292 feature points, only 53% is found conforming to the epipolar geometry. Thus, we detect nearly<sup>5</sup> 21% ( $1332/3292 \times 53\%$ ) true correspondences out of a total of 3292 feature points.

**Table 1.** Results by the BM and the CIFM for IP1 and IP2.

	BM-IP1	CIFM-IP1	BM-IP2	CIFM-IP2
Total fps	3292	3292	693	693
Detected fps	1332	1609	153	371
OutOfDetcd	53%	97%	54%	97%
OutOfTotal	21%	47%	12%	52%

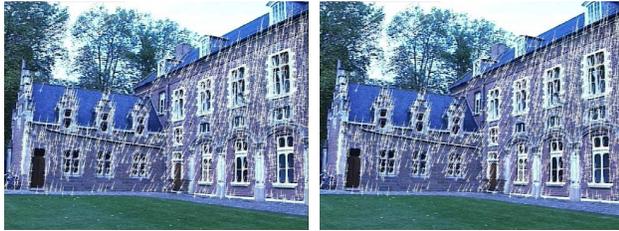
Fig. 4(c) and Fig. 4(d) show the correspondences obtained using the CIFM on IP1 before and after the outlier removal. From the figures, only few spurious correspondences are observed. From Table 1, for the CIFM-IP1, among the 1609 correspondences detected out of 3292 feature points, 97% are found conforming to the epipolar geometry. Thus, we nearly detected 47% true correspondences out of 3292 feature points.

Visual inspection of rows one and two of Fig. 4 also demonstrates that the CIFM significantly outperforms the BM in terms of both the number of the detected feature correspondence and the quality of the detected correspondences. We have applied the CIFM to

<sup>5</sup>Obviously, not all correspondences that comply to the epipolar geometry are true correspondences.



(a) BM-IP1 before outlier removal. (b) BM-IP1 after outlier removal.



(c) CIFM-IP1 before outlier removal. (d) CIFM-IP1 after outlier removal.



(e) BM-IP2 before outlier removal. (f) BM-IP2 after outlier removal.



(g) CIFM-IP2 before outlier removal. (h) CIFM-IP2 after outlier removal.

**Fig. 4.** Correspondences by the block matching (BM) technique and the proposed contrast-invariant feature matching (CIFM) algorithm on image pair 1 (IP1) and image pair 2 (IP2); the correspondences are illustrated by the displacement vectors superimposed on the first image of an image pair; outliers are removed using the epipolar constraint.

many other image pairs from the *castle* sequence. All results show the same observation.

Our second experiment is on an image pair (IP2) showing an evident change of the brightness. The two images were taken at the same time. However, the contrast of the two images differs significantly because the images contain different portions of the bright sky, causing different internal camera parameters.

Rows three and four of Fig. 4 show the results obtained by the BM and the CIFM on IP2, respectively, where we see that the CIFM

obtains much better results than the BM. From Table 1, for the BM-IP2, among the 153 correspondences detected out of 693 feature points, 54% are found conforming to the epipolar geometry. Thus, nearly 12% of true correspondences are detected out of a total of 693 feature points. In contrast, for the CIFM-IP2, 97% of the 371 correspondences detected out of 693 feature points conform to the epipolar geometry and thus nearly 52% true correspondences are detected.

As seen from Table 1, the CIFM is invariant to the image contrast. For the first image pair showing a small contrast difference, true correspondences are found for 47% of the total feature points. For the second image pair with evident contrast difference, the percentage of the true correspondences is 52%. The percentage keeps at a constant level regardless of the change of the contrast. In comparison, the percentage for the BM decreases from 21% for the first image pair to 12% for the second pair. Both are significantly lower than the percentages obtained by the CIFM. The reasons of the contrast invariance of the CIFM are two fold. First, the Harris corner detector is known to be robust to illumination change. Second, the CIFM relies only on the scene structure for feature matching.

## 5. CONCLUSIONS

In this paper, we have proposed a novel method for matching feature points between two images. The proposed method uses the geometry constraint alone for feature matching and is invariant to changes of the image contrast. The proposed technique has been applied to detect the feature correspondence for many image pairs. The experimental results show that the proposed method is: 1) invariant to the contrast, 2) able to obtain feature correspondences of much higher quality than block matching, 3) able to obtain a higher number of feature correspondences than block matching.

## 6. REFERENCES

- [1] Harris C. and Stephens M., "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conf.*, 1988, pp. 147–151.
- [2] Olga Veksler, "Semi-dense stereo correspondence with dense features," in *Proc. IEEE Workshop on Stereo and Multi-Baseline Vision*, 2001, pp. 149–157.
- [3] Joao Maciel and Joao P. Costeira, "A global solution to sparse correspondence problems," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 187–199, Feb 2003.
- [4] Abhijit S. Ogale and Yiannis Aloimonos, "Robust contrast invariant stereo correspondence," in *Proc. IEEE Int. Conf. Robotics and Automation*, April 2005, pp. 819–824.
- [5] G. Scott and H. Longuet-Higgins, "An algorithm for associating the features of two images," in *Proc. of the Royal Soc, London*, 1991, vol. B–244, pp. 21–26.
- [6] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–393, 1981.
- [7] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, "Visual modeling with a hand-held camera," *Int. Journal of Computer Vision*, vol. 59, no. 3, pp. 207–232, 2004.