The Eighth International Conference on Intelligent Technologies (InTech'07)
Sydney, Australia, 12-14 December 2007

Statistical Decision Based Gray-level Image Feature Matching

N. M. Kwok, M. G. Carmichael and Q. P. Ha
Faculty of Engineering
University of Technology, Sydney
Broadway NSW 2007, Australia
Email: nmkwok@eng.uts.edu.au
marc.g.carmichael@student.uts.edu.au
quangha@eng.uts.edu.au

K. C. Tan
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576, Republic of Singapore
Email: eletankc@nus.edu.sg

***Abstract*— The choices of gray-level image feature matching criteria and window sizes are considered. Features, as pixels enclosed in windows around interest points, are matched by operating on statistical measures where a match is declared in accordance to the result of a decision process. The decision error should be minimized as well as a small size window used in order to minimize the computational complexity. In this work, the maximum discrimination criteria is applied in the decision for a guaranteed bound on the matching error. The window size is determined by applying the principle of nonparametric sign-test, that satisfies the requirement for an adequate representation of the probability distribution function needed in conducting a match. The effectiveness of the proposed approach is verified by matching features in image sequences captured by a mobile robot, to the features detected and stored in the first frame.**

I. Introduction

Image and video processing techniques have found wide applications in areas such as medical imaging [1], surveillance [2], mobile robotics [3] and many others. These techniques, to name a few, include noise filtering, contrast enhancement, object segmentation, object recognition and feature matching. Here, an image feature is treated as a collection of neighbourhood pixels, forming a window, and possessing salient characteristics. Particularly, image feature matching or similarity checking is anticipated to play a critical role in the effectiveness of image processing in the above mentioned applications.

In practice, the use of image processing methods in mobile robotics for navigation [4] and localization [5] has been very attractive. It is also becoming popular with the employment of cameras as sensors for their low costs, low power consumptions and reduced payloads on mobile platforms. There are application examples in the context of localizing a mobile robot while producing a map of its operating environment [6] [7] and the references therein. The former work makes use of scale-invariant keypoints [8], characterized by the corresponding image descriptors, for their improved matching robustness against viewpoint changes. On the other hand, the latter work uses the gradients of the features in order to track them across sequential frames. Essentially, these methods require the extraction and tracking of salient image features across the captured image frame sequences so as to complete a task assigned to the robot. However, irrespective of the description or representation of the features [9], it is a mandatory requirement to track features by checking their similarities with respect to those obtained in previous frames or templates stored in a database as references.

Image feature matching can be viewed as image retrieval or registration processes as discussed in the surveys [10] and [11], where various techniques of feature matching had been reviewed. They range from area-based approaches by using correlation and frequency spectrums to that of feature-based methods using spatial relationships and invariance descriptors. Alternatively, statistical distance measures are also widely applied in checking feature similarities. In this regard, measures encompass those from the Bhattacharyya distance to that of the Euclidean distance [12] are available with a reducing order of complexity. Furthermore, their performances have been compared in an empirical study [13] together with other statistical measures. There are also feature matching tests particularly targeted at corner features, for example in [14] their performances were evaluated. In general, these approaches were developed with the focus on the use of statistical distances as a measure of image feature matching. A critical limitation in the distance based tests, however, is the need for a probability density function in describing the features. In view of this, a non-parametric test for similarity was proposed in [15]. Other methodologies adopted in feature matchings include those of information theoretic approach [16] which makes use of the entropy of pixels within an image window. An attempt using the multi-resolution distance measure was reported in [17] where sub-windows are formed by reducing the image sizes consecutively. There was also a similarity check by focusing on the correlation, changes in contrast and average intensity between features [18] as an alternative to statistical tests.

It is interesting to note that most approaches cited above require a specification of a window for a feature, hence, leads to the need for the choice of a proper window size. However, this issue has only been considered case-by-case in many reported applications. Furthermore, most image similarity checks are devoted to a comparison against thresholds where it is non-trivial to determine their optimum values. Consequently, this results in the lack of an assurance of the test quality, in particular, the probability of error incurred in declaring a match of similar images features.

In this work, the matching of gray-level image features is cast as a statistical decision problem. The problem scenario consists of matching features, e.g., corners, extracted from a sequence of images to that extracted from the first frame as references stored in a database. The differences in the gray-level intensities between two features from different frames are used as the test statistic. Histograms or probability distribution functions (pdf)[1] are constructed from these samples of image differences where the number of samples is determined by the size of the window embedding the feature. By invoking the Neyman-Pearson criteria with a user specified bound on the false-alarm rate, a decision of matching is declared. While proceeding in this manner, the principle of nonparametric sign test is adopted to determine the number of samples required, hence the window size, to provide a representative pdf to be used in the decision process. Image sequences captured from a monochrome camera mounted on a mobile robot navigating in an indoor environment are then used to verify the developed method.

The rest of the paper is organized as follows. In Section II, image feature matching measures are briefly reviewed and the motivation for this work is formulated. The matching procedure and the determination of window size are developed in Section III. Experimental results are given in Section IV to verify the proposed approach. Finally, a conclusion is drawn in Section V.

II. Feature Matching

The matching of image features or checking for their similarities has been tackled by a number of methods [10], [13]. They operate on the gray-level pixels within a pre-specified window while metrics or distances D are derived and then compared to some threshold γ in order to declare a match, i.e. $D \gtrless \gamma$. These metrics can be broadly classified into histogram-based and pixel-based measures. Metrics in the former category require the construction of a histogram while the latter use the intensities of the pixels directly. Typical and popular metrics are briefly presented below with an investigation into their limitations that motivate the present work.

A. Pixel-based Metrics

Consider two image features $\mathbf{X}_1$ and $\mathbf{X}_2$, each has an assumed window size of $n \times n$ containing $m = n^2$ pixels giving $\mathbf{X}_1 = \{X_1^1, \cdots, X_1^m\}$, for example.

[1]A histogram is the counts of occurrences of a particular gray-level while a distribution is a normalized histogram such that the area under the distribution is unity. These terminologies will be used interchangeably unless ambiguities arise in the context.

1) Mahattan distance: This distance simply measures the sum of absolute difference between the feature pixels which is equivalent to the $L_1 - norm$.

$$D = \sum_{k=1}^{m} |X_1^k - X_2^k|. \quad (1)$$

2) Euclidean distance: This is the conventional measure based on the multi-dimensional geometric distance, alternatively, it is the $L_2 - norm$.

$$D = \sqrt{\sum_{k=1}^{m} (X_1^k - X_2^k)^2}. \quad (2)$$

3) Canberra distance: The summation of the ratio between the difference and sum of pixel intensities is used as the metric in this test.

$$D = \sum_{k=1}^{m} \left\{ \frac{|X_1^k - X_2^k|}{|X_1^k + X_2^k|} \right\}. \quad (3)$$

4) Mahalanobis distance: The concept of normalized statistical distance between distributions, assumed Gaussian, is employed in this metric.

$$D = \sqrt{(\mathbf{X}_1 - \mathbf{X}_2)^T \Sigma^{-1} (\mathbf{X}_1 - \mathbf{X}_2)}, \quad (4)$$

where Σ is the covariance of vectorized features $(\mathbf{X}_1, \mathbf{X}_2)$ and the superscript T denotes a vector transpose.

B. Histogram-based Metrics

Consider an image feature $\mathbf{X}_1$, as in the pixel-based approaches, containing L gray-levels, e.g., $L = 256$. A histogram is constructed from counting the occurrence (frequency) of each gray-level such that $\mathcal{H} = \{H(0), \cdots, H(L-1)\}$, $H(i) \in [0, m]$ and $i = 0, \cdots, L-1$.

1) Histogram difference: This is a simple check against the absolute differences of the elements in the two histograms,

$$D = \sum_{i=0}^{L} |H_1(i) - H_2(i)|. \quad (5)$$

2) Histogram intersection: This metric attempts to award higher scores to features having similar histograms. That is,

$$D = \frac{\sum_{i=0}^{L} \min\{H_1(i) - H_2(i)\}}{\sum_{i=0}^{L} H_2(i)}. \quad (6)$$

3) χ^2 test: This metric operates on testing the difference between the two histograms in the χ^2 sense.

$$D = \sum_{i=0}^{L} \frac{(H_1(i) - H_2(i))^2}{H_1(i) + H_2(i)}. \quad (7)$$

4) Kolmogorov Smironov test: The maximum difference between the cumulative histograms is used as the test metric, that is

$$D = \max_i \{|H_1^c(i) - H_2^c(i)|\}, \quad (8)$$

where the superscript c stands for the cumulation of the histogram elements.

C. Limitations of Distance Metrics

Metrics making use of pixel intensities and histograms are justified by their own merits according to the problem domain and their effectiveness have been verified in many successful applications. However, there are several critical issues that need to be considered. First, it is difficult to determine the proper threshold, γ, which can guarantee a bounded error in declaring a match. It is also not being addressed, in most cases, on the choice of an optimal window size, $m = n^2$, which is to be related to the matching error. In the following, it is attempted to formulate a proper window size and to propose a matching method on the basis of satisfying a user-specified matching error limit.

III. Matching Threshold and Window Size

Matchings of image features are required before some assigned tasks, e.g., navigating a mobile robot using a vision sensor, can be accomplished. Accordingly, the declaration of a match between two image features is cast as a decision problem. By invoking decision theories [19], it is revealed that there are two kinds of errors incurred during a decision process. First, an error occurs when the features are different but declared as matched, i.e., a *false-alarm*, α. On the other hand, a declaration is not made when the features are actually similar, hence, producing a *miss-detection*, β, where the *power*[2] of the test is given by $1 - \beta$.

A. Matching Threshold

Consider the scenario where two features, e.g., corners [14], are to be checked for their similarity. Assume that a window size of $m = n^2$ pixels, has been determined and pixels are extracted into two arrays, $\mathcal{I}$ and $\mathcal{J}$. Their absolute differences are obtained from

$$\mathcal{E}_{uv} = |\mathcal{I}_{uv} - \mathcal{J}_{uv}|, \tag{9}$$

where $uv = 1, \cdots, m$ is the pixel index in the window.

The problem now becomes a decision to declare a match if the differences are *small* in order to obtain a low false-alarm rate. Moreover, it is also required to accept a match when the differences are *large* for a bounded miss-detection error rate. In order to achieve these two objectives, a probability density function (pdf) describing the pixel differences is required to calculate the false-alarm and miss-detection errors, α and β.

In the ideal case, let the two features be exactly matched when all the pixel differences are zero. This phenomenon can be visualized as a histogram, on the pixel differences, which contains an impulse at the first bin corresponding to the zero difference, i.e.,

$$H(0) = m, \; H(1 \leq i \leq L-1) = 0. \tag{10}$$

On the other hand, features not matched will have peaks in the histograms departing from the first bin, e.g.,

$$H(0) < \min_i \{H(i)\}, \; i \neq 0. \tag{11}$$

[2]Precisely, α is the probability of making a false-alarm while β is the probability of incurring a miss-detection, and the power indicates the probability in declaring a match when the features are similar.

As an example, consider the instance when comparing a database feature in Fig. 1(a) to a candidate feature, Fig. 1(b), in the current frame where the camera has been displaced. False-alarms can be treated as being caused by interferences coming from other features which are indicated by the dotted-squares in Fig. 1(b).

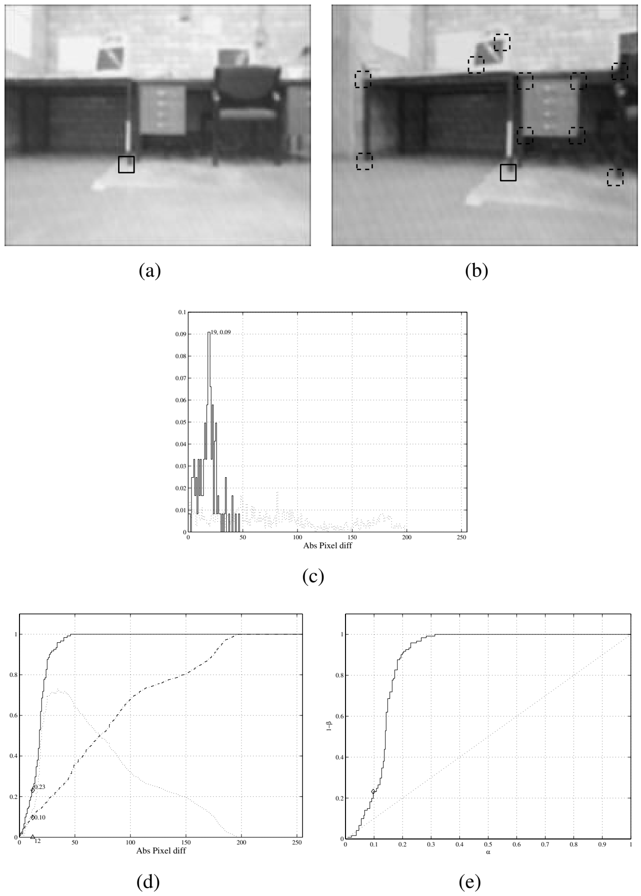
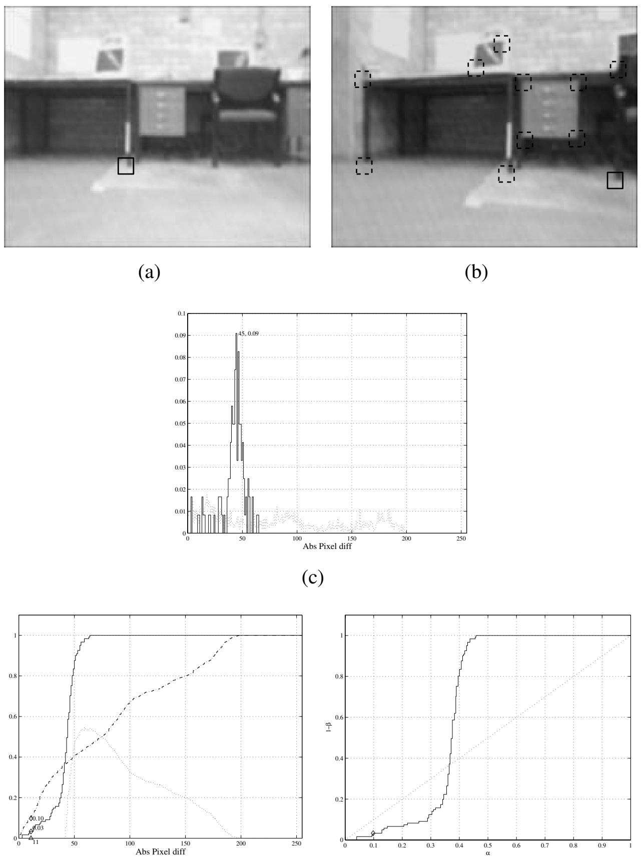


Fig. 1. Image features to be matched, matched case. (a) database feature (solid square), (b) candidate feature to be compared (solid square), interfering features (dotted square). Test statistics for a matched case. (c) histograms (solid - from feature to be compared, dotted - from other features), (d) cumulative histogram (upper - power, lower - significance and differences - dotted line), (e) ROC curve, $\diamond$ is the decision point.

The pixel differences (between the database and candidate feature) then form the histogram as plotted in the solid-line in Fig. 1(c) with a peak at the 19^{th} difference and having the distribution at 0.09. Another histogram, derived from all the pixel differences due to the interfering features is shown with the dotted-line. The power of the test is given by the cumulative sum of the distribution of the candidate pixel difference (upper curve in Fig. 1(d)) while the significance is illustrated by the lower curve. At a significance of 0.10, the power is 0.23 at the 12^{nd} pixel difference. The separation between the power and significance is indicated by the dotted-line in the figure. This curve is an illustration of the discriminative capability of the test and a positive discrimination is derived in this case. An assessment of the test is also available from the receiver operating characteristic (ROC) curve as depicted in Fig. 1(e). The decision point is above the diagonal and confirms a discriminative test. In this example, it is assured that the decision for a match is made with a test that bounds a false-alarm rate within 0.10. The power of the test is obtained at 0.23 and is larger than the false-alarm rate.

Another example of feature matching is shown in Fig. 2(a) and 2(b). In this case, the candidate feature is intentionally selected to be different from the database feature as an illustrative counter-example. The candidate feature is located on the lower-right hand corner of the image.



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

Fig. 2. Image frames to be matched, no-match case. (a) database feature (solid square), (b) input features (feature in solid square to be compared, features in dotted square are interferences). Test statistics for no-match case. (c) histograms (solid - from feature to be compared, dotted - from other features), (d) cumulative histogram (upper - power, lower - significance and differences - dotted line), (e) ROC curve, $\diamond$ is the decision point.

The histogram, Fig. 2(c), illustrates that the pixel differences at small values are sparse and, consequently, demonstrates a low similarity between the database and candidate feature. The power and significance curves are shown in Fig. 2(d). It is noted that at the 0.10 false-alarm level the power is almost zero (at 0.03), hence, there is no discrimination and a declaration for a match cannot be made in this case. Furthermore, the decision point shown in the ROC curve, Fig. 2(e) is below the diagonal line and no discrimination is available from the test.

The above development has demonstrated that the matching decision can be made on the basis of the decision error probability calculated from the distributions of the candidate and interfering features. However, it is important to ensure that the distribution is capable of representing the statistics of the pixel differences. Intuitively, in order to obtain a properly represented distribution, a large number of samples (pixel differences) is required. Specifically, this is related to the size of the window that embeds the feature to be compared.

B. Window Size

An exact representation of the feature image difference distribution is deemed impracticable unless an infinitive number of samples is drawn from the differences. This observation then leads to the need for determining the number of samples such that the representation error is bounded.

The strategy adopted here, to determine the window size, is to minimize the difference between two distributions (one as a general distribution, see (11), and the other is the available histogram under test). Sample points are drawn to perform tests in the sense of *goodness-of-fit*, while the test samples required and the window size are related to the *significance* of the test of goodness.

Assume that a true distribution $\mathcal{T}\{\epsilon\}$ describing the feature differences is available, where $\epsilon = \epsilon(uv)$ is an element in the difference $\mathcal{E}$, see (9). Here, the available distribution is constructed from the feature differences, while the ideal distribution is obtained from all-zeros elements (for exact feature match). Then a set of S sample pairs are drawn from the true distribution and the available feature differences distribution simultaneously. That is

$$\mathcal{X} = \{\varsigma_1, \cdots, \varsigma_S\},\ \varsigma_i = \{\mathcal{E}_{uv}, \mathcal{T}_{uv}\},\ i = 1, \cdots, S, \qquad (12)$$

where each ς_i contains a feature difference value from the image and another value from the true distribution.

As sample pairs are sequentially drawn (index i increases), evidence is being collected to reject the claim that the available distribution corresponds to the ideal match. After a sufficient number of samples are drawn, the claim can be rejected with some specified confidence. By invoking the non-parametric sign test technique, see Appendix, a positive sign is assigned to a sample pair if their difference is below some threshold (e.g., 1 gray-level). The number of samples required can be derived by relating the number of signs to the significance ε of the test by

$$\sum_{n=\kappa+1}^{S} \left(\begin{array}{c} S \\ n \end{array} \right) \left(\frac{1}{2}\right)^S = \varepsilon. \qquad (13)$$

The needed number of samples κ is then converted to the window size as the minimum number $m = n^2$ (for $n = 3, \cdots$ and odd) such that $m > \kappa$.

From tables of cumulative binomial distribution [19], the error incurred from 8 signs out of 9 samples is 0.002. As a further illustration, this significance can be reduced to 0.0003 if 15 signs are available from 16 samples. That is, the test significance reduces as more samples are used. With a 256 gray-level image, there are 256 bins in the error histogram. In order to cover the histogram, we choose a window size of $n = 25$ giving $m = n^2 = 625$ pixels. In this case, if the bins are filled correctly by more than half of the error samples e.g. $342 > 625/2$, then the error in determining a proper pdf representation becomes $1 - \sum_{i=342}^{625} \mathcal{B}(i, 625, 0.5) = 0.0082$ with an error less than 1%.

IV. Results

Tests of feature matching are conducted on a sequence of images obtained from a camera mounted on-board a mobile robot, while moving in a anti-clockwise circular trajectory in an indoor laboratory environment. The mobile robot travels on a flat terrain such that the images captured are different from each other in a sequentially shifting manner.

The well proven Harris detector is used to extract the features which represent corners of the furniture in the laboratory. In order to verify the proposed method, features in subsequent frames are matched to those obtained in the first frame. Notice that in this test scenario, features disappear from the camera field-of-view after about 20 frames (due to the characteristic of the camera lens) and re-appears when the robot returns to a similar spot where the feature was last seen. This arrangement specifically tests for the ability to match features at different viewpoints.

The sequence captured from the on-board camera are 200×150 (width$\times$height) 256-level black-and-white images. Snapshots of the image sequence are shown in Fig. 3. A total of 500 frames are stored and the matching is performed off-line on the MATLAB® platform. Examples are shown for the early stage: frames 2 to 5, the mid-travel phase: frames 200 to 212 and late-stage: frames 413 to 423.

In the figures, the first reference frame is depicted on the bottom-right in each sub-figure shown. When point features are matched, lines from the top-left plots are drawn to their corresponding points at the bottom-right (first frame) plots. In the early stage (Fig. 3(a), 3(b)), images are captured while the mobile is at similar distances to the features as stored in the first frame. Matchings are satisfactory while the robot is turned slightly in the anti-clockwise rotation. In the mid-travel stage (Fig. 3(c), 3(d)), the robot is displaced at a longer distance from its location when the first image was taken. This is evident as the chairs appear on the right hand side that were not previously seen. Here, notwithstanding the interference from the newly observed features, matching to the original features are not degraded. At the late travel stage (Fig. 3(e), 3(f)), the robot returns to a location approximately the same as the mid-travel phase. Satisfactory matching results are still maintained as illustrated. In summary, the proposed image feature matching approach is robust to plane rotation and longitudinal translation of the camera.

V. Conclusion

Matching of image features are studied in this paper. Difficulties and limitations in determining threshold-based matching methods, including pixel-based and histogram-based matrices, are revealed. This paper proposed a hypothetical testing approach that is able to relax the need to find such thresholds while the matching performance is related to a guaranteed decision error bound. Furthermore, the choice of window size that contains the feature is discussed. Matching examples are given with satisfactory performances using the proposed methodology and its effectiveness is further verified by matching features from a sequence of images captured by a mobile robot.

Acknowledgment

This work is supported by UTS partnership grant RM#2006000848 and the Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

References

[1] I. Frosio, G. ferrigno and N. A. Borghese, "Enhancing digital cephalic radiography with mixture models and local gamma correction," *IEEE Trans. on Mdeical Imaging*, vol. 25, no. 1, pp. 113-121, 2006.

[2] R. Cucchiara, C. Grana, A. Prati and R. Vezzana, "Computer vision system for in-house vedio survelliance," *IEE Proc. Vis. Image Signal Processing*, vol. 152, no. 2, pp. 242-249, 2005.

[3] P. Saeedi, P. D. Lawrence and D. G. Lowe, "Vision-based 3-D trajectory tracking for unknown environments," *IEEE Trans. on Robotics*, vol. 22, no. 1, pp. 119-136, 2006.

[4] D. Murray and J. J. Litle, "Using real-time stereo vision for mobile robot navigation," *Autonomous Robots*, vol. 8, pp. 161-171, 2000.

[5] H. Hajjdiab and R. Laganière, "Vision-based multi-robot simultaneous localization and mapping," *Proc. First Canadian Conf. on Computer and Robot Vision*, pp. 155-162, 2004.

[6] S. Se, D. Lowe and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *The Intl. J. of Robotics Research*, vol. 21, no. 8, pp. 735-758, 2002.

[7] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera," *Proc. Ninth IEEE Intl. Conf. on Computer Vision*, pp. 1403-1410, 2003.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Intl. J. of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[9] F. Fraundorder, H. Bischof and S. Ober, "Natural, salient image patches for robot localization," *Proc. 17th Intl. Conf. on Patern Recognition*, pp. 881-884, 2004.

[10] S. Antani, R. Kasturi and R. Jain, "A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and videos," *Pattern Recognition*, vol. 35, pp. 945-965, 2002.

[11] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, pp. 977-1000, 2003.

[12] N. Vasconcelos and A. Lippman, "A unifying view of image similarity," *Proc. 15th Intl. Conf. on Pattern Recognition*, pp. 38-41, 2000.

[13] M. Kokáre, B. N. Chatterji and P. K. Biswas, "Comparison of similarity metrices for texture image retrieval," *Proc. Conf. on Convergent Technologies for Asia-Pacific Region, TENCON 2003*, pp. 571-575, 2003.

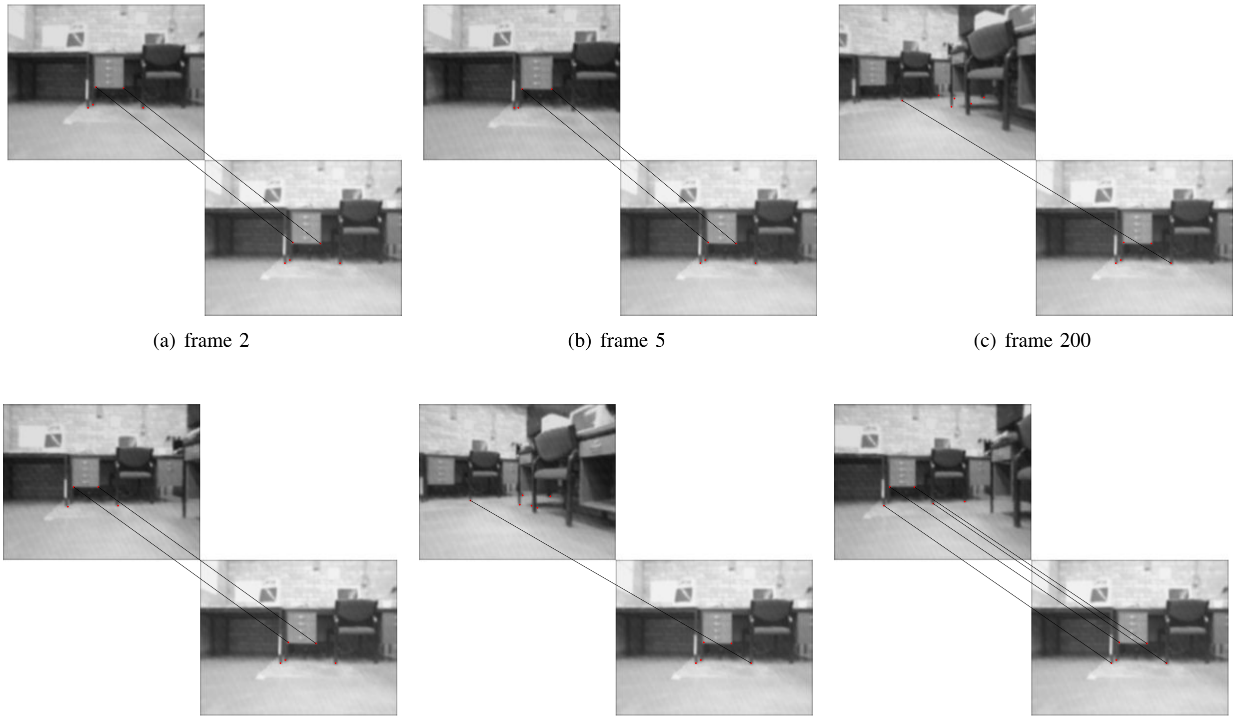
[14] P. Tissainayagam and D. Suter, "Assessing the performance of corner detectors for point features tracking applications," *Image and Vision Computing*, vol. 22, pp. 663-679, 2004.

[15] J. Puzicha, T. Hofmann and J. M. Buhmann, "Non-parametric similarity measures for unsupervised texture segmentation and image retrieval," *Proc. IEEE Comp. Society Conf. on Computer Vision and Pattern Recognition*, pp. 267-272, 1997.

[16] G. Fritz, L. Paletta and H. Bischof, "Object recognition using local information content," *Proc. 17th Intl. Conf. on Pattern Recognition*, pp. 15-18, 2004.

[17] P. Juffs, E. Beggs and F. Deravi, "A multiresolution distance measure for images," *IEEE Signal Processing Letters*, vol. 5, no. 6, pp. 138-140, 1998.

[18] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81-84, 2002.



(a) frame 2 (b) frame 5 (c) frame 200

(d) frame 212 (e) frame 413 (f) frame 423

Fig. 3. Snapshots of image matching results at different images frames.

[19] J. L. Melsa and D. L. Cohn, *Decision and estimation theory*, New York: McGraw-Hill, 1978.

APPENDIX

The sign test [19] is a hypothesis testing method that operates independently of the test statistic distribution. The basic principle is developed from testing the median of a distribution. The test problem may be described by

$$\begin{aligned} m_1 &: \; P\{z_s > 0 | m_1\} = 1/2 \\ m_2 &: \; P\{z_s > 0 | m_2\} = p > 1/2, \end{aligned} \tag{14}$$

where z_s is the sample drawn, and (m_1, m_2) are the messages transmitted from the corresponding hypothesis.

Since the test makes use only the sign of the difference between the hypothesis, then a change of variable gives

$$w_i = u(z_i) = \begin{cases} 1, & z_i > 0 \\ 0, & z_i \leq 0 \end{cases}. \tag{15}$$

The problem statement can be re-written as

$$P\{w_i | m_1\} = \begin{cases} 1/2, & w_i = 0 \\ 1/2, & w_i = 1 \end{cases} \tag{16}$$

and

$$P\{w_i | m_2\} = \begin{cases} 1-p, & w_i = 0 \\ p, & w_i = 1 \end{cases}. \tag{17}$$

In terms of w_i, a likelihood ratio can be written as

$$\wedge(w) = \frac{p^{\sum_{i=1}^{S} w_i}(1-p)^{S - \sum_{i=1}^{S} w_i}}{(\frac{1}{2})^S}. \tag{18}$$

Putting $S^* = \sum_{i=1}^{S} w_i$ as the number of positive observations, the likelihood ratio test becomes

$$\wedge(w) = [2(1-p)]^S \left(\frac{p}{1-p}\right)^{s^*} \gtrless \kappa^*, \tag{19}$$

and taking logarithm to base $p/(1-p)$, then

$$S^* \gtrless \kappa^* - S \log_{\frac{p}{1-p}} [2(1-p)] = \kappa. \tag{20}$$

The threshold is determined by setting the false-alarm probability below, ε, for instance. Since S^* is the sum of S Bernoulli random variables, it obeys the binomial distribution. If message m_1 is true, the binomial distribution has parameters S and $1/2$, so that

$$\begin{aligned} P\{S^* = n | m_1\} &= \binom{S}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{S-n} \\ &= \binom{S}{n} \left(\frac{1}{2}\right)^S. \end{aligned} \tag{21}$$

The false-alarm is

$$P\{S^* > \kappa | m_1\} = \sum_{n=\kappa+1}^{S} \binom{S}{n} \left(\frac{1}{2}\right)^S = \varepsilon. \tag{22}$$

Hence, for a given number of samples S, the smallest value of κ is sought such that the decision error is bounded by the specified ε.