Research Article

Robust Multisensor Image Matching Using Bayesian Estimated Mutual Information

Lurong Shen, Xinsheng Huang, Yuzhuang Yan, Yongbin Zheng, and Wanying Xu

College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, Hunan 410073, China

Correspondence should be addressed to Lurong Shen; shenlurong09@gmail.com

Received 24 May 2013; Accepted 26 August 2013

Academic Editor: Marco Perez-Cisneros

Copyright © 2013 Lurong Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mutual information (MI) has been widely used in multisensor image matching, but it may lead to mismatch among images with messy background. However, additional prior information can be of great help in improving the matching performance. In this paper, a robust Bayesian estimated mutual information, named as BMI, for multisensor image matching is proposed. This method has been implemented by utilizing the gradient prior information, in which the prior is estimated by the kernel density estimate (KDE) method, and the likelihood is modeled according to the distance of orientations. To further improve the robustness, we restrict the matching within the regions where the corresponding pixels of template image are salient enough. Experiments on several groups of multisensor images show that the proposed method outperforms the standard MI in robustness and accuracy and is similar with Plum's method. However, our computation is far more cost saving.

1. Introduction

The matching of multisensor images is an important preprocessing step for many applications, such as navigation, data-fusion, and visualization tasks. Due to different physical characteristics of different sensors, the relationship between the intensities of the matching pixels is often complex. Visual features present in one sensor image may not appear in other ones, and vice versa; image contrast may differ from each other in the same regions; multiple intensities in one image may be mapped to a single intensity in other images, and vice versa. The uncorrelated intensities make the normal intensity-based matching methods ineffective; therefore, the matching of multisensor image becomes challenging in computer vision.

Previous works on multisensor image matching can be generally categorized into two classes [1]. The first class is to find some invariant representations among multisensor images, such as point feature [2], edge feature [3], contour feature [4] and edge orientation maps [5], then the problem can be solved by utilizing the traditional monosensor matching methods. Nevertheless, the feature extraction has been

regarded as one of the most difficult problems in image processing, especially in multisensor cases, as the feature present in one image may not appear in other ones. Additionally, these methods usually omit most of the image details, which may lead to a wrong matching.

The second class is to use an invariant similarity measure of images. These methods are conducted directly in the intensity domain, and they have been successfully used in a variety of multisensor registrations. Within these methods, the images are treated as combinations of different sets. Each set stands for a separate gray level, assuming that the same component of the same object is consistent in one image, and the pixels of the same component with intensities near x in one image correspond to the pixels with intensities near y in the other image. In other words, there may be no pixel-to-pixel correspondence in intensity, but there do exist set-to-set correspondences between two images; namely, ignoring the intensity property, the topology of the corresponding sets must be similar between two images. The corresponding relationship can be well shown by the joint intensity scatter plot (JISP) or the joint histogram (JH) [6], and by measuring their dispersion one can quantify the matching degree. One

FIGURE 1: A mismatching example of MI. (a) An infrared template image, (b) a visible reference image and the matching result, and (c) the matching result of our method. There are many background textures around the target (car) of the infrared image, but only a few in the VS image, which decrease the statistical relationship and lead to a mismatching when using the NMI method. The matching result is shown by superimposing the Sobel edges of the template image directly on the reference image.

of the most successful measures is the entropy. The lower the entropy is, the more tightly clustered the scatter plot will be and, hence, the closer the matching will be. Mutual information (MI) [7, 8] or normalized mutual information (NMI) [9] has been considered as the state-of-the-art matching frameworks of the entropy-based methods.

Nevertheless, the robustness of the MI-based matching methods is questionable. A mismatched result of the NMI between an infrared (IR) and a visible (VS) image is shown in Figure 1, in which statistical relationship between the images is decreased by the messy background. In addition, a “Grey stripe” registration also demonstrates the pitfall of NMI in [10].

Recently, many approaches have been proposed to improve the MI by incorporating additional spatial information. Studholme et al. [11] introduce a prior of labeled connecting regions as an additional information channel to the MI. Guo and Lu [12] adopt the gradient vector flow (GVF) to enhance the original intensity images before using the MI, which improves the success rate. Kim et al. [13] propose a new method based on a 3D joint histogram (JH), which integrates the intensity information and the edge orientation information. The method provides more robustness. Lee et al. [14] extend [13] by using a modified GVF to extend the capture range of gradient-based MI. Wang and Tian [15] encode the images into 32 bins according to the gradient magnitude and orientation and then compute the MI directly on the encoded field, which improves the computation efficiency of the MI. Liu et al. [16] improved this method with an adaptive combination strategy of an intensity-based MI and a coded gradient-based MI. Fan et al. [17] propose a method that combines the spatial information through a feature-based selection mechanism, such as the Harris corner, and the JH is increased to 4D. Loeckx et al. [18] propose conditional mutual information (cMI) as a new similarity for image registration; cMI starts from a 3D joint histogram incorporating, besides the intensity dimensions, also a spatial dimension expressing the location of the joint intensity pair. Pluim et al. [19] modify the MI-based method by multiplying

the MI with a term of the gradient. The gradient term not only seeks to align locations of high gradient magnitude but also aims for a similar orientation of the gradients at these locations. This method is referred to as GNMI in this paper.

These improvements for the MI-based methods strongly depend on the invariant characteristics mentioned in the first class, for example, the gradient information of magnitude and orientation, as well as the corner features. Actually, since they take full advantages of the two classes, we classify them as the third class and regard this additional information as prior information in this paper.

It should be pointed out that some methods which introduce the prior into MI are problematic. For example, in [15], the pixels are encoded according to their gradient magnitude and orientation. The intensities in the standard MI are substituted with gradient codings. It is known that the gradient orientation appears to be parallel to each other in multisensor images, and they must be distributed along the diagonal of JH in the true matching. However, as the codings function similarly as the intensities of the standard MI in JH, and MI measures the dispersion of JH only, this method still gives a high correlation even though the compact bins are located far away from the diagonal of JH. Moreover, the gradient magnitudes participating in the coding are not so reliable, as the magnitudes vary a lot in multisensor images [13, 14]. Some better methods are presented in [13, 16, 19], in which they treat the measure of gradient vectors between multisensor images as a weight of the MI. However, they suffer from expensive computation.

In this paper, a new similarity measure that incorporates the prior information into MI by the Bayesian method is proposed. For simplicity, it is named as Bayesian MI (BMI) here. BMI is deduced by substituting classical probabilities in MI with the Bayesian probability, such as the prior, the likelihood, and the posterior. For example, the classical conditional probability can be regarded as a likelihood probability or a posterior probability in the Bayesian framework and through which we can introduce some subjective experiences in the estimation. To avoid the statistic of JH, the BMI is

defined directly on each pixel pair, and the images can then be measured by the mean BMI of the valid matching pixel pairs.

Then, an implementation of the BMI is proposed based on the prior information of gradient. Since the magnitude information between multisensor images is not so reliable, we only concern whether their magnitudes are salient or not and treat the salient ones as the saliency. Generally, saliency of image contributes more in matching [20, 21], so we restrict the pixel pairs of matching within a mask where the pixels of the template image are salient.

In the implementation, the orientation of each pixel is defined firstly. The saliency pixels are defined as the absolute value of their orientations, while the nonsaliency ones are defined as zero. Secondly, the prior distribution is estimated based on the kernel density estimate (KDE) method, and the likelihood between orientations is modeled by the Gaussian function according to their distance. The closer the orientation is, the more similar the pixel pair will be. Finally, the sequential statistic BMI based on the gradient information is modeled as well.

At the end of this paper, experiments are conducted on six groups of images which are captured by different sensors, and the results suggest that the proposed method is more robust and more accurate than the standard NMI and is similar to Plum's method [19]. Moreover, our computation is far more cost saving.

The rest of this paper is outlined as follows. Section 2 briefly introduces the standard MI, and the proposed BMI is deduced by Bayesian formula. Based on the gradient prior information, the application of the BMI is shown in Section 3. In Section 4, some experimental results and discussions are presented. Finally, conclusions are drawn in Section 5.

2. Proposed Method

2.1. Mutual Information. The MI method is of general significance because it makes few assumptions on image intensities. It assumes neither linear correlation nor even functional correlation but only statistical dependence [10].

Given two images, A and B , one can define their joint probability (JP) $p_{AB}(a, b)$ by simply normalizing the JH. Let $p_A(a)$ and $p_B(b)$ denote the corresponding probabilities, which can be obtained directly from $p_{AB}(a, b)$. The MI between A and B based on Shannon entropy is given by

$$\begin{aligned} \text{MI}(A, B) &= H(A) + H(B) - H(A, B) \\ &= \sum_{a,b} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_A(a) p_B(b)}. \end{aligned} \quad (1)$$

Apart from (1), MI can also be described in many other forms, as summarized by [22]:

$$\begin{aligned} \text{MI}(A, B) &= H(A) - H(A|B) \\ &= \sum_{a,b} p_{AB}(a, b) \log \frac{p_{A|B}(a|b)}{p_A(a)}, \end{aligned} \quad (2)$$

where $H(A)$ and $H(B)$ are the Shannon entropy of images A and B , which indicate the amount of uncertainty about A and B , respectively. $H(A|B)$ and $H(B|A)$ denote the conditional entropy, indicating the uncertainty left in A or B when B or A is known. $H(A, B)$ is the joint entropy, which indicates the total uncertainty of A and B . When b is known, the conditional probability of a is noted as $p_{A|B}(a, b)$; it can be easily deduced from their JP by the classical probability theory.

In fact, $-\log(a)$ can be regarded as the self-information of a , and the entropy $H(a)$ represents its mean. Similarly, the mutual information between a and b can be defined as

$$\text{MI}(a, b) = \log \frac{p_{AB}(a, b)}{p_A(a) p_B(b)}, \quad (3)$$

and $\text{MI}(A, B)$ can be deemed as the mathematical expectation of $\text{MI}(a, b)$.

2.2. Bayesian Mutual Information. BMI is deduced directly from the MI by regarding the conditional probability $p_{A|B}(a|b)$ in (2) as a posterior probability, which can be estimated according to the prior information based on the Bayesian method.

Let $p_A(a)$ be the prior probability; then, according to the Bayesian formula, we have

$$\hat{p}_{A|B}(a|b) = \frac{\hat{p}_{B|A}(b|a) p_A(a)}{\hat{p}_B(b)} = \frac{\hat{p}_{B|A}(b|a) p_A(a)}{\sum_a \hat{p}_{B|A}(b|a) p_A(a)}, \quad (4)$$

where $\hat{p}_{B|A}(b|a)$ and $\hat{p}_B(b)$ denote the likelihood probability and evidence probability, respectively. $\hat{p}_B(b)$ can be estimated through the joint of prior and likelihood by the total probability formula. Here, $\hat{p}$ is used for distinguishing the Bayesian probabilities from the classical probability p . Using (2) and (4), the BMI can be defined as

$$\text{BMI}(A, B) = \sum_{a,b} p_{AB}(a, b) \log \frac{\hat{p}_{B|A}(b|a)}{\sum_a \hat{p}_{B|A}(b|a) p_A(a)}. \quad (5)$$

Similar to the MI, (5) needs the statistic of JP beforehand. For convenience, we refine (5) into a sequential statistic form. As $p_{AB}(a, b) = (1/N) \sum_{x,y} \delta(a_{x,y} - a) \cdot \delta(b_{x,y} - b)$, where N is the total pixel number in the matching, $\delta(x)$ is the Kronecker function, $a_{x,y}$ and $b_{x,y}$ are the pixel values of A and B at (x, y) , respectively. Thus, (5) can be rewritten as

$$\begin{aligned} \text{BMI}(A, B) &= \frac{1}{N} \sum_{x,y} \text{BMI}(a_{x,y}, b_{x,y}) \\ &= \frac{1}{N} \sum_{x,y} \log \frac{\hat{p}_{B|A}(b_{x,y}|a_{x,y})}{\hat{p}_B(b_{x,y})} \\ &= \frac{1}{N} \sum_{x,y} \log \frac{\hat{p}_{B|A}(b_{x,y}|a_{x,y})}{\sum_{a_{x,y}} \hat{p}_{B|A}(b_{x,y}|a_{x,y}) p_A(a_{x,y})}, \end{aligned} \quad (6)$$

where $\text{BMI}(a_{x,y}, b_{x,y})$ denotes the BMI of the pixel pair $(a_{x,y}, b_{x,y})$. Obviously, a higher BMI $\text{BMI}(a_{x,y}, b_{x,y})$ can be achieved from higher likelihood and lower evidence. Equation (6) is applicable regardless of whether a and b are discrete or continuous, while (5) is for the discrete form only.

BMI can be comprehended favourably using the communication theory when regarding $a_{x,y}$, $\hat{p}_{B|A}(b_{x,y} | a_{x,y})$ and $b_{x,y}$ as the message source, the channel, and the sink; respectively, BMI is just the evaluation of the correct rate after transmission [23].

It should be noticed that $p_A(a_{x,y})$, $\hat{p}_{B|A}(b_{x,y} | a_{x,y})$, $\hat{p}_B(b)$, and their logarithmic values can be calculated offline; moreover, as $\log(xy) = \log(x) + \log(y)$, the cost of our computation can be remarkably saved by utilizing the lookup-table (LUT) technique.

3. Implementation

In this section, we are about to implement the proposed BMI based on the prior information of the gradient. Actually, BMI is universal as long as the prior and the likelihood can be modeled, such as the intensity and the local statistic between monosensor matching, the gradient orientation and magnitude, and the edges but not the intensity of multisensor images.

As mentioned in the introduction, the saliency includes most of the information [24]; thus, it would contribute more in matching. In [21], Qin et al. extract the salient vector for each pixel using the principal axis analysis (PAA). The more similar the corresponding vectors are, the higher weight the pixel pairs will be in JP. In [20], Kim et al. extract the salient regions from expanding the edges; then the MI is calculated on those regions in which the images are both salient. This method enhances the statistical correlation between the IR and VS images.

In this study, we simply define the saliency as those pixels whose gradient magnitude are higher than a threshold, which is correlated to the mean of the full image. Different from Kim's scheme, the proposed BMI is computed only on those pixel pairs in which the template image is salient, regardless of the reference image.

Let R and T denote the reference image and the template image, respectively, and R_{uv} denotes the subimage cropped from the top-left corner (u, v) of R , which has the same size as T . Both of the two images are filtered by a 5×5 Gaussian window to suppress the noise.

The realization of the BMI includes the gradient orientation mapping, the prior probability estimation, and the likelihood probability modeling

It should be noted that, in the following implementation, it is assumed that there is only translation between two images, so are the experiments. This assumption does not affect manifesting the effect of our method.

3.1. Gradient Orientation Mapping. Let $\theta_{x,y}^T$ and $\theta_{x,y}^{R_{uv}}$ be the gradient orientation maps of T and R_{uv} , and (x, y) is the

coordinate of pixel. Taking $\theta_{x,y}^T$ for an example, it is defined as

$$\theta_{x,y}^T = \begin{cases} \text{angle}(V_{xy}), & \text{if } (\nabla T_y > 0 \cap G_{xy} > t), \\ p, & \text{if } (\nabla T_y = 0 \cap G_{xy} > t), \\ 0, & \text{if } G_{xy} < t, \end{cases} \quad (7)$$

$$V_{xy} = \text{sgn}(\nabla T_y) \begin{bmatrix} \nabla T_x \\ \nabla T_y \end{bmatrix}, \quad G_{xy} = |V_{xy}|,$$

where $\text{angle}(x)$ is a function which returns the phase of vector x , G_{xy} is its magnitude, and $\text{sgn}(\cdot)$ is the sign function, which reverses the vectors in the 3rd and 4th quadrant and limits the $\theta_{x,y}^T$ to $[0, \pi]$. τ is the threshold which separates the gradient of salient structures from noises, and

$$\tau = \frac{a}{N} \sum_{x,y} G_{xy}, \quad (8)$$

where a is an empirical percentage of the mean magnitude, N is the total pixel number of I . Practically, $a \in [0.8, 1.2]$. Here, the fact that gradient orientations are parallel to each other between multisensor images is utilized. In other words, their orientations are either the same or opposite. When the gradient magnitude is small, $\theta_{x,y}^T$ is set to zero, and the pixels with horizontal orientation are defined as π .

Let $\Delta\theta$ be the resolution angle; then we can get the discrete form of (7)

$$c_{x,y}^T = \text{round} \left(\frac{\theta_{x,y}^T + 0.5}{\Delta\theta} \right). \quad (9)$$

In this paper, the pixels of $\theta_{x,y}^T = 0$ are named as zero-code pixels, and those of $\theta_{x,y}^T \neq 0$ are named as nonzero-code pixels, and those nonzero-code pixels are defined as saliency. The saliency of T forms the mask in which the matching is conducted, and we denote the set of the corresponding coordinations (x, y) as V .

3.2. Prior Probability Density Estimation. Normally, the prior probability can be defined by the histogram of $C_{x,y}^T$. However, when the size of the template image is small (small sample size problem), the histogram will be incapable to character the probability precisely by the classical probability theory, which requires as many samples as possible. For this problem, kernel density estimate (KDE), or Parzen window method [25, 26], is a good solution. KDE has been widely used for nonparametric distribution estimation [27]. Therefore, the probability density of $\theta_{x,y}^T > 0$ can be written as

$$f_T(\theta) = \frac{1}{C_{h1}} \sum_{(x,y) \in V} K_1 \left(\left\| \frac{\theta - \theta_{x,y}^T}{h_1} \right\|^2 \right), \quad (10)$$

where $K_1(\cdot)$ is the kernel function and h_1 corresponds to the bandwidth. The larger h_1 is, the smoother $f_T(\theta)$ will be. Constant C_{h1} makes $\int f_T(\theta) d\theta = 1$.

FIGURE 2: Valid pixel pairs in matching are sorted into two classes: nonzero-code to nonzero-code and nonzero-code to zero-code, which are marked with "1" and "2" respectively. The saliency (nonzero-code) is marked by diagonal lines.

Actually, the KDE method distributes each $\theta_{x,y}^T$ to its neighbors by a bounded $K_1(\cdot)$ with h_1 and makes $f_T(\theta)$ more accurate and more robust than the discrete one. Frequently used kernel functions include Uniform, Triangle, Epanechnikov, and Normal [28]. In this study, the Normal Gaussian kernel function is employed as kernel function.

3.3. Likelihood Probability Density Model. Likelihood is a subjective probability in our work. It is a connection between the prior and the evidence. In this section, we sort the valid pixel pairs under the mask into two classes: nonzero-code to nonzero-code matching and nonzero-code to zero-code matching, as shown in Figure 2. Their likelihoods will be modeled, respectively.

As for the first class, its likelihood can be modeled according to their distance between $\theta_{x,y}^T$ and $\theta_{x,y}^{R_{uv}}$. It is clear that the closer the orientation is, the higher the likelihood will be. Taking two random orientations, θ_r and θ_t , as an example, their likelihood probability density can be modeled as

$$\hat{f}_{R_{uv}|T}(\theta_r | \theta_t) = \frac{1}{C_{h_2}} K_2\left(\frac{\angle(\theta_r, \theta_t)}{h_2}\right), \theta_r, \theta_t > 0, \quad (11)$$

where $K_2(\cdot)$ is a hat-like function which is high in the middle but low on both sides, for example, the Gaussian function. h_2 defines its bandwidth. $\angle(\theta_r, \theta_t)$ is the intersection angle between the two orientations, and $0 \leq \angle(\theta_r, \theta_t) \leq \pi/2$, which is calculated by

$$\angle(\theta_r, \theta_t) = \min\{|\theta_r - \theta_t - \pi|, |\theta_r - \theta_t|, |\theta_r - \theta_t + \pi|\}. \quad (12)$$

C_{h_2} is a constant which ensures $\hat{f}_{R_{uv}|T}(\theta_r | \theta_t)$ satisfy the boundary condition of $\int \hat{f}_{R_{uv}|T}(\theta_r | \theta_t) d\theta_r = 1$.

Figure 3 shows an example of $\hat{f}_{R_{uv}|T}(\theta_r | \theta_t)$ when $h_2 = 12/\pi$. It indicates that the bigger the intersection angle is, the lower the likelihood will be, and it approaches to 0 as $\angle(\theta_r, \theta_t)$ tend to $\pi/2$. Combining (10) and (11), the evidence probability density can be obtained:

$$\hat{f}_{R_{uv}}(\theta) = \int f_T(\theta_t) \cdot \hat{f}_{R_{uv}|T}(\theta | \theta_t) d\theta_t. \quad (13)$$

FIGURE 3: The likelihood probability density $\hat{f}_{R_{uv}|T}(\theta_r | \theta_t)$ versus the intersection angle $\angle(\theta_r, \theta_t)$ based on the Gaussian function $h_2 = 12/\pi$.

FIGURE 4: Curves of $f_T(\theta)$ and $\hat{f}_{R_{uv}}(\theta)$ under different h_2 . Apparently, $\hat{f}_{R_{uv}}(\theta)$ is a smoothed result of $f_T(\theta)$. The bigger h_2 is, the smoother $\hat{f}_{R_{uv}}(\theta)$ will be.

The continuous function, $\hat{f}_{R_{uv}}(\theta)$, is just a smoothed form of $f_T(\theta)$ by using the function K_2 , and they will be overlapped when $h_2 \leq \pi/180$, as shown in Figure 4.

With regard to the nonzero-code to zero-code matching, it can be treated as mismatching, and its likelihood can be defined as 0. Finally, we have

$$\begin{aligned} \hat{f}_{R_{uv}|T}(\theta_r | \theta_t) &= \begin{cases} \frac{1}{C_{h_2}} K_2\left(\frac{\angle(\theta_r, \theta_t)}{h_2}\right), & \text{if } (\theta_r, \theta_t > 0), \\ 0, & \text{if } (\theta_r = 0, \theta_t > 0). \end{cases} \end{aligned} \quad (14)$$

3.4. Practical Issues. Now, the BMI of R_{uv} and T can be derived by substituting $a_{x,y}$, $b_{x,y}$, A , and B in (6) with $\theta_{x,y}^T$, $\theta_{x,y}^{R_{uv}}$, T , and R_{uv} , respectively. First, some issues should be noticed.

- (1) BMI will be defined as 0 if the likelihood equals 0, such as the second class.

FIGURE 5: Ten groups of IR reference images.

FIGURE 6: Ten groups of VS reference images.

- (2) We can see that there do exist some pixel pairs that satisfy $\hat{f}_{R_{uv}|T}(\theta_{x,y}^{R_{uv}} | \theta_{x,y}^T) < \hat{f}_{R_{uv}}(\theta_{x,y}^{R_{uv}})$, which results in a negative BMI. Sometimes, the mean BMI may be negative, too, especially in those mismatched locations. But it never occurs in the standard MI [22]. In this study, the pixel pairs with negative BMI are regarded as mismatched, and their BMIs are limited to zero, so that they are consistent with the second class.

Finally, the BMI-based matching formula between R_{uv} and T can be derived as

$$\begin{aligned} & \text{BMI}(T, R_{uv}) \\ &= \frac{1}{N} \\ & \times \sum_{(x,y) \in V} \begin{cases} \log \frac{\hat{f}_{R_{uv}|T}(\theta_{x,y}^{R_{uv}} | \theta_{x,y}^T)}{\hat{f}_{R_{uv}}(\theta_{x,y}^{R_{uv}})}, & \text{if } \left(\frac{\hat{f}_{R_{uv}|T}}{\hat{f}_{R_{uv}}} > 1, \theta_{x,y}^{R_{uv}} \neq 0 \right), \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (15)$$

The true matching of the template image T can be obtained by maximizing the $\text{BMI}(T, R_{uv})$ through out the interval (u, v) of the reference image R .

4. Experimental Validation

To evaluate the performance of the proposed method, firstly, ten pairs of IR and VS images are used in the test as shown in Figures 5 and 6, respectively. The IR images are captured by a long-wave IR camera, whose sizes are 108×168 , and they are taken as the template image; the VS images are captured by a CCD camera, whose sizes are 256×256 , and they are taken as the reference images. All the images are captured on a sunny afternoon, with only some random translation transform between them. As the ground truth is unknown for the sets of real images, it is estimated by manually matching in Photoshop. We also test these four methods on two pairs of image scrapped from [29], as shown in Figures 7 and 8.

The matching results obtained from the proposed method are compared with the NMI, GNMI [19], and GCMI methods [15]. These methods are chosen because they have been widely used in multisensor image matching, and they provide relatively good matching performance. In the NMI and GNMI methods, the densities are estimated by JH with a fixed number of $64 * 64$ bins for both images. The intensity values are linearly mapped into these bins. The parameters used in these methods are $h_1 = 10\pi/180/3$, $h_2 = 90\pi/180/3$, and $a = 0.8$, the matching results of our method are shown in Figure 10.

FIGURE 7: (Group 11) A pair of downtown images captured at night and in daytime.

FIGURE 8: (Group 12) A pair of images of Mauna Loa.

The comparison results of the NMI, GNMI, GCMI, and our method are shown in Figure 9. A global searching strategy was applied to get a full view of the correlation curve. The figures indicate the following.

- (1) The BMI outperforms the NMI significantly. It can be seen that (1) the difference between the highest and the 2nd higher peak, which measures the interference suppressing ability, is more salient in BMI than in NMI through all the tests. In particular, the NMI mismatches in the 8th, 9th, and 11st test, while the BMI

is still applicable. (2) The highest peak of BMI, which measures the position accuracy, is much sharper than that of NMI. The matching results are shown in Table 1, from which we can see that all the methods have matching errors, but our method is more accurate.

- (2) Because GNMI, GCMI, and BMI adopt gradient information as measure, the performances of our method are similar to the GNMI and GCMI; they all have similar matching accuracy and the salient highest peak. However, as discussed previously, the

(a)

FIGURE 9: Continued.

FIGURE 9: Matching curves of the BMI (the 4th row) compared with the NMI (the 1st row), GNMI (the 2nd row), and GCMi (the 3rd row).

FIGURE 10: Matching results of the BMI.

TABLE 1: Comparison of the position accuracy.

Group	Ground truth	NMI/error	GNMI/error	GCMI/error	BMI/error
1	(72, 44)	(72, 37)/(0, 7)	(73, 44)/(1, 1)	(73, 45)/(1, 1)	(73, 44)/(1, 0)
2	(66, 16)	(65, 15)/(1, 1)	(65, 15)/(1, 1)	(65, 15)/(1, 1)	(65, 15)/(1, 1)
3	(38, 57)	(39, 57)/(1, 0)	(39, 57)/(1, 0)	(39, 57)/(1, 0)	(39, 57)/(1, 0)
4	(104, 43)	(105, 43)/(1, 0)	(105, 41)/(1, 2)	(105, 42)/(1, 1)	(105, 42)/(1, 1)
5	(49, 45)	(49, 44)/(0, 1)	(49, 44)/(0, 1)	(49, 44)/(0, 1)	(49, 44)/(0, 1)
6	(99, 14)	(103, 13)/(4, 1)	(99, 13)/(0, 1)	(99, 13)/(0, 1)	(99, 14)/(0, 0)
7	(69, 28)	(70, 26)/(1, 2)	(70, 26)/(1, 2)	(70, 25)/(1, 3)	(70, 26)/(1, 2)
8	(107, 55)	(142, 54)/(35, 1)	(107, 55)/(0, 0)	(107, 55)/(0, 0)	(107, 55)/(0, 0)
9	(118, 38)	(143, 78)/(35, 40)	(119, 38)/(1, 0)	(118, 37)/(0, 1)	(119, 37)/(1, 1)
10	(97, 41)	(97, 41)/(0, 0)	(96, 41)/(1, 0)	(96, 41)/(1, 0)	(96, 41)/(1, 0)
11	(192, 396)	(33, 398)/(159, 2)	(192, 396)/(0, 0)	(191, 396)/(1, 0)	(192, 396)/(0, 0)
12	(55, 306)	(55, 306)/(0, 0)	(55, 306)/(0, 0)	(55, 306)/(0, 0)	(55, 306)/(0, 0)
Mean error	(—/—)	(19.7, 5.4)	(0.6, 0.7)	(0.7, 0.8)	(0.6, 0.5)

The bold data means mismatching pairs.

computation of our method is remarkably cost saving when the LUT technique is used. In the MATLAB environment, the cost ratio of the four methods is about 1 : 2 : 3 : 4 (BMI : GCMI : NMI : GNMI), and the proportion can be further improved when using other program tools, such as the C and C++ language.

Additional experiments have also been done to test the robustness according to the parameter settings, including a , h_1 , and h_2 , which refer to (10), (11), and (14), respectively. The results are listed below.

- (1) When $a < 0.8$, the noise increases and it makes the matching curve of BMI not so smooth, but the main peak of the curve remains unchanged. When $a > 0.8$, the salient regions shrink, and the position accuracy decreases.
- (2) The change of h_1 has no substantial influence on $\hat{f}_{R_{uv}}(\theta)$, as long as $h_1 < h_2$.
- (3) h_2 determines the cover range of the likelihood, and the higher h_2 is, the smoother the curve will be and, hence, the more robust the BMI will be.

5. Conclusions

In this paper, a robust Bayesian estimated mutual information method for multisensor images matching is proposed. Our method is universal as long as the prior and the likelihood can be modeled, and we have implemented it by using the prior information of gradient, in which the prior probability density is estimated by the KDE method, and the likelihoods are modeled, respectively, according to the class of the matching pixel pair. Our method is compared with the standard NMI and GNMI by experimenting on some multisensor images, and the results show that (1) our method provides higher robustness and position accuracy than NMI; (2) as the keys of BMI can be precalculated offline, our method is remarkably cost saving; (3) the parameters of the BMI can be set easily, which facilitate its application.

Further studies include improving the matching accuracy by suppressing the speckle noise and refining the saliency mask. Moreover, to enhance its robustness, additional information can be introduced according to the practical conditions.

References

- [1] M. Irani and P. Anandan, "Robust multi-sensor image alignment," in *Proceedings of the 6th IEEE International Conference on Computer Vision*, pp. 959–966, January 1998.
- [2] J. Chen and J. Tian, "Real-time multi-modal rigid registration based on a novel symmetric-SIFT descriptor," *Progress in Natural Science*, vol. 19, no. 5, pp. 643–651, 2009.
- [3] L. Huang and Z. Li, "Feature-based image registration using the shape context," *International Journal of Remote Sensing*, vol. 31, no. 8, pp. 2169–2177, 2010.
- [4] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Medical Image Analysis*, vol. 1, no. 2, pp. 151–161, 1996.
- [5] Y. Keller and A. Averbuch, "Multisensor image registration via implicit similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 794–801, 2006.
- [6] J. Orchard and R. Mann, "Registering a multisensor ensemble of images," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1236–1247, 2010.
- [7] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.
- [8] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Proceedings of the 5th International Conference on Computer Vision*, pp. 16–23, June 1995.
- [9] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.
- [10] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal

- image registration,” in *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '98)*, vol. 1496 of *Lecture Notes in Computer Science*, pp. 1115–1124, Springer, Cambridge, Mass, USA, October 1998.
- [11] C. Studholme, D. L. G. Hill, and D. J. Hawkes, “Incorporating connected region labelling into automated image registration using mutual information,” in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '96)*, pp. 23–31, June 1996.
- [12] Y. Guo and C.-C. Lu, “Multi-modality image registration using mutual information based on gradient vector flow,” in *Proceedings of the 18th IEEE International Conference on Pattern Recognition (ICPR '06)*, vol. 3, pp. 697–700, August 2006.
- [13] Y. S. Kim, J. H. Lee, and J. B. Ra, “Multi-sensor image registration based on intensity and edge orientation information,” *Pattern Recognition*, vol. 41, no. 11, pp. 3356–3365, 2008.
- [14] J. H. Lee, Y. S. Kim, D. Lee, D.-G. Kang, and J. B. Ra, “Robust CCD and IR image registration using gradient-based statistical information,” *IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 347–350, 2010.
- [15] X. Wang and J. Tian, “Image registration based on maximization of gradient code mutual information,” *Image Analysis and Stereology*, vol. 24, pp. 1–7, 2005.
- [16] J. Liu, J. Tian, and Y. Dai, “Multi-modal medical image registration based on adaptive combination of intensity and gradient field mutual information,” *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 1429–1432, 2006.
- [17] X. Fan, H. Rhody, and E. Saber, “A spatial-feature-enhanced MMI algorithm for multimodal airborne image registration,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2580–2589, 2010.
- [18] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, “Nonrigid image registration using conditional mutual information,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 19–29, 2010.
- [19] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, “Image registration by maximization of combined mutual information and gradient information,” *IEEE Transactions on Medical Imaging*, vol. 19, no. 8, pp. 809–814, 2000.
- [20] K. S. Kim, J. H. Lee, and J. B. Ra, “Robust multi-sensor image registration by enhancing statistical correlation,” in *Proceedings of the 7th International Conference on Information Fusion*, pp. 380–386, July 2005.
- [21] B. Qin, Z. Gu, X. Sun, and Y. Lv, “Registration of images with outliers using joint saliency map,” *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 91–94, 2010.
- [22] J. P. W. Pluim, J. B. A. A. Maintz, and M. A. Viergever, “Mutual-information-based registration of medical images: a survey,” *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.
- [23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, NJ, USA, 2nd edition, 2006.
- [24] D. Mahapatra and Y. Sun, “Rigid registration of renal perfusion images using a neurobiology-based visual saliency model,” *Eurasip Journal on Image and Video Processing*, vol. 2010, Article ID 195640, 16 pages, 2010.
- [25] E. Parzen, “On estimation of a probability density function and mode,” *The Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.
- [26] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: criteria of Max-Dependency, Max-Relevance, and Min-Redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
- [27] P. Thévenaz and M. Unser, “Optimization of mutual information for multiresolution image registration,” *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2083–2099, 2000.
- [28] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [29] G. Yang, C. V. Stewart, M. Sofka, and C.-L. Tsai, “Registration of challenging image pairs: initialization, estimation, and decision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1973–1989, 2007.