AN AUTOMATIC IMAGE-MAP ALIGNMENT ALGORITHM BASED ON MUTUAL INFORMATION AND HILBERT SCAN

Li Tian, Sei-ichiro Kamata

Graduate School of Info., Pro. and Sys., Waseda University
2-7, Hibikino, Wakamatsu-ku, 808-0135, Kitakyushu, Japan
phone: + (81) 93-692-5219, email: tianli@ruri.waseda.jp,kam@waseda.jp

ABSTRACT

An algorithm for automatic image-map alignment problem using a new similarity measure named Edge-Based Code Mutual Information (EBCMI) and Hilbert scan is presented in this study. Because image and map are very different in their representations, the normal Mutual Information (MI) using the intensity in traditional alignment method may result in misalignment. To solve the problem, codes which are robust to the differences between the image-map pairs are constructed and Mutual Information of the codes is computed as the similarity measure for the alignment. We convert the 3-D transformation search space in alignment to a 1-D search space sequence by using 3-D Hilbert Scan. A new search strategy is also proposed on the 1-D search space sequence. The experimental results show that the proposed EBCMI outperformed the normal MI and some other similarity measures and the proposed search strategy gives flexibility between efficiency and accuracy for automatic image-map alignment task.

1. INTRODUCTION

Image-map alignment is a special multimodal image alignment problem. It is important for updating Geographic Information System (GIS) via revising the digital map by using the aerial image. Some similarity measures for multimodal images such as medical images have been proposed over the years [1]. Mutual Information (MI) is currently one of the most popular measures of them because it is based on global information of images and does not require the landmarks or features for alignment. It is an automatic and intensity-based measure. Its flexibility and accuracy have been shown for aligning multimodal medical images [2]. Gradient information has also been combined with MI for medical images in work [3, 4].

However, because the image and the map are quite different from each other in their representations where both the intensities and constructions are changed greatly, image-map alignment is different from the multimodal medical image alignment. Figure 1 shows an aerial image and a digital map of the same scene in the urban area. The major differences in their representations can be concluded as following:

- Modelling the objects such as bridge, railway, and buildings in the image to corresponding signs for the map results in differences in their intensities and constructions;
- Some symbols and characters are added to the map for noting the names of places such as streets and buildings. These symbols can be viewed as noise.
- Several objects visible in the image such as moving cars and trees are not in the map.

Figure 1: The image and the map of the same scene.

- Because the image and the map are usually taken in different time, changes may occur between them.

Therefore, the normal MI based on intensity is not suitable for this task, and similarity measures which can cope with the differences in representations for automatic image-map alignment are highly desired.

Many methods for automatic image-map alignment have been developed in the recent years. A fully automatic system for the alignment of satellite image data with vector GIS data is presented in [5] where polygonal objects are extracted from the image and the GIS data set. However, Ground Control Points (GCPs) are required and extracted from vector GIS data in this system while it is difficult to extract GCPs from the digital map data in our case. Two similarity measures based on the Hausdorff distance are also proposed in [6, 7] to solve the automatic alignment. In addition, an algorithm to extract common feature from image and map using the diffusion process [8] has also been proposed. These methods are useful in images containing large smooth regions like in medical images, but often fail in images without them as the image-map pair of urban area shown in Figure 1.

In this study, we propose a new similarity measure for automatic image-map alignment named Edge-Based Code Mutual Information (EBCMI) to solve the difficulties. First, codes based on the edges of both the image and the map are constructed by analyzing the differences in their representation. Then, MI function based on the codes, rather than the intensity, is computed as the similarity measure for alignment. In this study, we use the Shannon measure of entropy [9]. An important property of the proposed codes, as well as EBCMI, is that it is robust to the differences between the image and map in representation.

On the other hand, we also propose a new search strategy based on 3-D Hilbert scan. Hilbert scan is a special scan and Figure 2 gives an example of a 3-D Hilbert scan. In recent

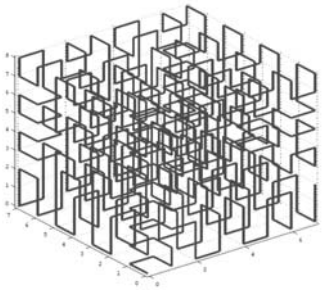


Figure 2: An example of a 3-D Hilbert scan.

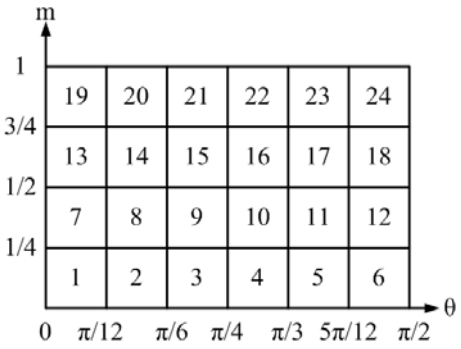


Figure 3: An example of the EBC.

years, it has been widely applied in image processing tasks such as image compression, clustering an image, pattern recognition and point pattern matching [12, 13, 14, 15, 16]. More details about 3-D Hilbert Scan for arbitrarily-sized arrays can be found in work [11]. We first convert the multi-dimensional search space (x,y translation and rotation) into 1-D sequence by using 3-D Hilbert scan, and a search strategy which is similar to gradient descent is proposed. We use several image-map pairs to evaluate our method in the experiment, the proposed EBCMI shows its superiorities to the normal MI and some other similarity measures, and the search strategy gives flexibility between efficiency and accuracy for automatic image-map alignment task.

The rest of this paper is organized as follows. Section 2 is about the proposed EBCMI. In Section 3, we give details of the search strategy based on 3-D Hilbert Scan. Experiments using the proposed algorithm and comparisons with other conventional methods for automatic image-map alignment are described in Section 4. We conclude this study in the last section.

2. SIMILARITY MEASURE

In this section, the proposed similarity measure EBCMI was first introduced in [17]. We will recall its definition at first. Then, we discuss the reasons why the EBCMI is robust to the differences between the image and the map in representation.

2.1 EBCMI

Given a discrete image $f(x,y)$ defined in 2D space, the horizontal and vertical gradients of it are presented as $f'_h(x,y)$ and $f'_v(x,y)$. Then a clipped magnitude $m(x,y)$ of the gradient vector is calculated as

$$m(x,y) = \rho(\sqrt{(f'_h(x,y))^2 + (f'_v(x,y))^2}) \tag{1}$$

where ρ is a clipping function defined as

$$\rho(x) = \begin{cases} x & \text{if} \quad x \geq t \\ 0 & \text{else} \end{cases} \tag{2}$$

with a predefined threshold t. And an absolute angle $\theta(x,y)$ of the gradient vector can be obtained by

$$\theta(x,y) = tan^{-1}(|\frac{f'_v(x,y)}{f'_h(x,y)}|) \tag{3}$$

where $|\cdot|$ is the absolute operation. The range of $m(x,y)$ is set to be $[0,1]$ for convenience, and $\theta(x,y)$ is thus in range $[0,\pi/2]$ by using the absolute operation.

The EBC is constructed from the clipped magnitudes and absolute angles of images obtained above. First, the obtained clipped magnitudes and absolute angles are divided into n_m and n_θ intervals by using two constants Δ_m and Δ_θ

$$n_m = \frac{1}{\Delta_m}, \qquad n_\theta = \frac{\pi}{2\Delta_\theta}. \tag{4}$$

The EBC is defined as

$$c(x,y) = \lfloor \frac{m(x,y)}{\Delta_m} \rfloor n_\theta + \lfloor \frac{\theta(x,y)}{\Delta_\theta} \rfloor + 1 \tag{5}$$

where $\lfloor x \rfloor$ presents the floor integer of x. This coding process is similar to the one in [4].

The EBC is supposed to be robust to the differences between the image and the map in representation and we will discuss it later. Figure 3 gives an example of the EBC when $\Delta_m = 1/4$ and $\Delta_\theta = \pi/12$. For coded image A and coded map B, the MI of them is

$$MI_{EBC}(A,B) = \sum_{a,b} p_{AB}(a,b) \log \frac{p_{AB}(a,b)}{p_A(a)p_B(b)}. \tag{6}$$

It can be written as the following equations

$$\begin{aligned} I(A,B) &= H(A) + H(B) - H(A,B) \\ &= H(A) - H(A|B) \\ &= H(B) - H(B|A) \end{aligned} \tag{7}$$

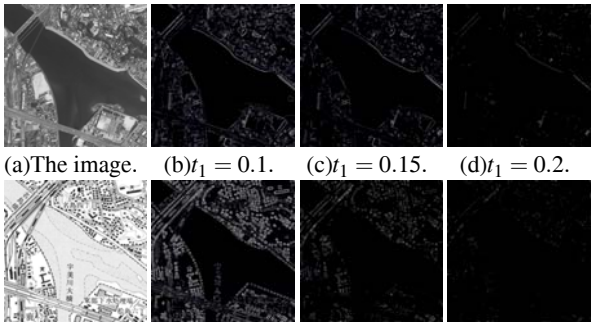
where $H(A)$ and $H(B)$ are the separate entropies of A and B, respectively. $H(A,B)$ is their joint entropy, and $H(A|B)$ and $H(B|A)$ are the conditional entropies of A given B and of B given A, respectively. The above entropies also can be written as

$$H(A) = -\sum_a p_A(a) \log p_A(a), \tag{8}$$

$$H(B) = -\sum_b p_B(b) \log p_B(b), \tag{9}$$

$$H(A|B) = -\sum_{a,b} p_{AB}(a,b) \log p_{A|B}(a|b) \tag{10}$$

according to the Shannon measure of entropy.



(a)The image. (b)$t_1 = 0.1$. (c)$t_1 = 0.15$. (d)$t_1 = 0.2$.

(e)The map. (f)$t_2 = 0.2$. (g)$t_2 = 0.3$. (h)$t_2 = 0.4$.

Figure 4: Clipped magnitude of images and maps by varying different t.

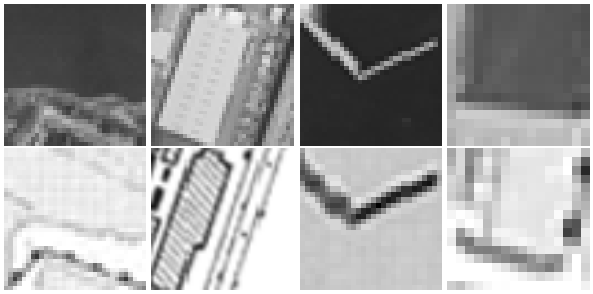


Figure 5: Some examples of modelled signs.

2.2 Properties of EBCMI

As mentioned previously, the image and the map are quite different from each other in their representations where the intensities and constructions are changed greatly. Here, we will analyze how do the two factors: the clipped magnitude and the absolute angle of gradient in EBC overcome the differences in their representations in detail.

First, we calculate the clipped magnitude of gradient instead of the normal one as in Eq. (2). This calculation is proposed for ignoring the small gradient in low contrast region. Since more intensity levels are given in the image than in the map, the clipping of low contrast region can reduce the differences in intensity levels between the image and map. However, too large t may cause the information lost of important edges with large magnitude of gradient. Figure 4 shows some illustrations of the clipped magnitude by varying different t for both the image and the map. Notice that the thresholds for the image and the map are different, usually, larger for the map.

Second, absolute angle of the gradient is used in the EBC as in Eq. (3). Objects such as bridge, railway, and buildings in the image are modelled to corresponding signs for the map. In many cases, objects in the image are modelled as inverse or near inverse signs for the map. Figure 5 shows some examples of the modelled signs. We can find that the signs in the map tend to be inverse in intensity when compared with the objects in the ie. Hence, the absolute angle of the gradient can reduce the effect of the 'inverse' changes.

3. SEARCH STRATEGY

In this section, we will explain how to obtain the optimal transformation parameters by using Hilbert scanning technique and analyze the computational complexity of the search strategy.

3.1 Proposed Method

In this study, since scaling transformation can be known as a prior factor from the scales of the image and the map, we only consider the translation and rotation including horizontal translation Δx, vertical translation Δy and rotation θ. The proposed alignment algorithm can be listed as the following:

1. Construct the 3-D parameter space coordinates (Δ x, Δ y, θ);
2. Sample transformation space to discrete points;
3. Use 3-D Hilbert scan to convert the 3-D parameter space into 1-D space;
4. Find the optimal parameters using the proposed search strategy, which will be discussed later, using EBCMI as the similarity measure.

We use Hilbert scan to convert the 3-D space into 1-D sequence as $H = \{H_1, H_2, ..., H_L\}$ where L is the total number of the points in the 3-D parameter space. Each point in the 1-D sequence corresponds to a transformation and has an EBCMI value. We should find a point whose EBCMI value is the largest. Of course, when the number of the points becomes large, it takes much time to do the search. Here we will show how to search more efficiently.

Some notions used in the search strategy are as follows:

C_i : the ith candidate point in current candidate set;

$S = \{C_i | i = 1, 2, 3..., N\}$: current candidate set including N candidates;

M_i: the middle point of the left or right adjacent intervals of C_i;

C_{max}: the candidate point with the largest EBCMI value in current candidate set;

G_{max}: the candidate point with the largest EBCMI value in whole search procedure;

$EBCMI(x)$: the EBCMI value of a candidate point x;

$T\%$: the selection ratio for keeping the candidate to the next search stage.

The strategy which is similar to gradient descent is as follows:

Step 1 Initialize G_{max} with a small number and equally divide the whole 1-D sequence into k parts using $k+1$ initial points with interval $\Delta d = \Delta d_0$.

Step 2 Form S with the $k+1$ initial points.

Step 3 For C_i, compare its two adjacent points $C_i + \Delta d$ and $C_i - \Delta d$. If $EBCMI(C_i + \Delta d) \geq EBCMI(C_i - \Delta d)$, $M_i = C_i + \Delta d/2$; otherwise, $M_i = C_i - \Delta d/2$;

Step 4 Compare M_i to C_i, if $EBCMI(M_i) \geq EBCMI(C_i)$, replace C_i with M_i;

Step 5 Repeat Steps 3 and 4 until all candidates are processed.

Step 6 Find C_{max} from S, if $C_{max} \geq G_{max}, G_{max} = C_{max}$;

Step 7 Rank all the candidate and delete duplicate candidate, then keep the top $T\%$ candidates to the next search stage;

Step 8 $\Delta d = \Delta d/2$. If $\Delta d \geq 1$, go to step 3); otherwise, stop and go to next step;

Step 9 Let point with G_{max} be the optimal solution.

Here, $T\%$ is to control the selection. When $T\%$ is smaller, fewer candidates will be selected. Thus, $T\%$ gives flexibility

Optimal (θ°, Δ x, Δ y)	Pair 1 (-3,-21,-23)	Pair 2 (0,-16,-17)	Pair 3 (-10,-52,-60)	Pair 4 (10,-52,-64)
EBCMI	(-1,-20,-22) *(2,1,1)*	(0,-11,-19) *(0,5,2)*	(-10,-50,-62) *(0,2,2)*	(10,-51,-59) *(0,1,5)*
Normal MI	(-1,-20,-22) *(2,1,1)*	(0,-11,-19) *(0,5,2)*	(0,-50,-60) *(10,2,0)*	(0,-51,-61) *(10,0,3)*
GCMI	(-1,-20,-22) *(2,1,1)*	(0,-11,-19) *(0,5,2)*	(-8,-50,-58) *(2,2,2)*	(9,-51,-57) *(1,1,7)*
VWMHD	(0,-21,-22) *(3,0,1)*	(0,-10,-18) *(0,6,1)*	(-9,-50,-50) *(1,2,10)*	(0,-52,-50) *(10,0,14)*

Table 1: Alignment results by using different measures (translation is in pixels and the numbers in the latter bracket indicate the errors).

between speed and accuracy in the search.

3.2 Computational Complexity

We can analyze the computational complexity of the proposed search strategy as in [13]. Two extreme cases may occur in the search process: first, all C_is are replaced by M_is in each stage and we only need to compute EBCMI value at M_is. If C_{ave} is the average number of candidates selected for each stage, the number of total search points in this case is

$$N_1 = k + 1 + C_{ave}\log_2(\Delta d_0); \tag{11}$$

Second, if no C_i is replaced by M_i in each stage and we have to compute EBCMI value at $C_i + \Delta d$ and $C_i - \Delta d$, which results in the number of search points becomes

$$N_2 = k + 1 + 2C_{ave}\log_2(\Delta d_0). \tag{12}$$

Here, Δd_0 can be written as

$$\Delta d_0 = \frac{L}{k}. \tag{13}$$

Thus, the average number of N_1 and N_2 is treated as the final estimation of computational complexity

$$N_{ave} = k + 1 + \frac{3}{2}C_{ave}\log_2(\frac{L}{k}). \tag{14}$$

Because the linear item k increases faster than the logarithmic item, the computational complexity essentially depends on the initial candidate number k. Selecting a proper k is important for the proposed strategy.

4. EXPERIMENTAL RESULTS

In this section, we design three experiments to show the performance of the proposed algorithm: the first one is to compare EBCMI with some other similarity measures in image-map alignment; the second one for tuning the parameter t in our algorithm; the last one is to demonstrate the flexibility of our search strategy.

In the first experiment, we use four image-map pairs as the test data as shown in Figure 6, where the image which is smaller than the map in each pair and the size of them is in range from 256×256 pixels to 325×325 pixels. Without loss of generality, we fix the image and only allow the map to transform. We compare the proposed ECBMI with the normal MI [2], the GCMI [4] and the VWMHD [7]. The optimal transformation parameters are obtained manually. The search range for rotation is $[-15^\circ, 15^\circ]$ centered at the optimal rotation with step 1°, and is $[-15, 15]$ centered at the optimal translation with step 1 pixel for both the x and y translations. We use the full search in this experiment. The

(a)Pair 1 (b)Pair 2 (c)Pair 3 (d)Pair 4

Figure 6: Image-map pairs used in the experiment.

(t_1,t_2)	(0.1,0.2)	(0.15,0.3)	(0.2,0.4)
(θ°, Δ x, Δ y)	(9,-51,-57)	(10,-51,-59)	(0,-51,-58)

Table 2: Alignment results varying different t for image-map pair 4.

alignment results using different measures are shown in Table 1. From the table, we can see that all the measures obtain similar results near the optimal transformation positions in image-map pair 1 and 2 with large smooth regions. However, it is easy to see that the proposed EBCMI gets better results than Normal MI and VWMHD in image-map pair 3 and 4 without large smooth regions. Moreover, EBCMI is slightly better than GCMI which is also based on edge code. We believe that the improvement is from the clipped magnitude and the absolute angle of gradient.

We also apply an alignment experiment when t changes in the clipping function using image-map pair 4. Table 2 shows the alignment results when t changes (t_1 for the image and t_2 for the map). The results show the effect of the clipping function in our EBCMI and an inappropriate t may degrade its performance greatly. From experiment using many image-map pairs, we find that $t_1 = 0.15$ and $t_2 = 0.3$ are the best choice in most cases.

Finally, we set up an experiment when T changes in the selection function using image-map pair 4. The search range for rotation is $[-9^\circ, 10^\circ]$ centered at the optimal rotation with step 1° and is $[-9, 10]$ centered at the optimal translation with step 1 pixel for both the x and y translations. Thus, the 3-D Hilbert used here is $20 \times 20 \times 20$. The initial number k here is 20. Table 3 shows the alignment results when T changes and the number of searched point $\#N$. We obtain the same results for three different Ts. This phenomenon indicates that by choosing proper number of initial candidates k and T, the search strategy can give flexibility between speed and accuracy in the search.

T	100	50	10
(θ°, Δ x, Δ y)	(3,-53,-65)	(3,-53,-65)	(3,-53,-65)
#N	173	26	12

Table 3: Alignment results and number of searched point #N by varying different T for image-map pair 4.

5. CONCLUSIONS AND FUTURE WORK

In this study, we propose a new similarity measure named Edge-Based Code Mutual Information (EBCMI) and a new search strategy for automatic image-map alignment task. EBCMI can overcome the differences in presentations between the image and the map and the new search strategy can give flexibility between accuracy and speed. The experimental results using the proposed similarity is better than using other conventional similarity measures and the new search strategy can balance the accuracy and efficiency.

The future studies will aim at automatically setting the parameters and removing the redundant information such as characters in the map. Some non-rigid deformation between the image and the map should also be considered.

6. ACKNOWLEDGEMENT

This work was supported at part by the JSPS Research Fellowships for Young Scientists.

REFERENCES

[1] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical Image Analysis*, Vol. 2, No. 1, pp. 1–36, 1998.

[2] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Medical Imaging*, Vol. 22, No. 8, pp. 986–1004, 2003.

[3] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Medical Imaging*, Vol. 19, No. 8, pp. 809–814, 2000.

[4] 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.

[5] H. Hild, "Automatic image-to-map-registration of remote sensing data," *Photogrammetric Week 2001*, pp. 13–23, 2001.

[6] L. Tian, S. Kamata, Y. Ueshige, and Y. Kuroki, "An automatic image-map registration algorithm using modified partial hausdorff distance, " in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Vol. 5,pp. 3534–3537, 2005.

[7] L. Tian and S. Kamata, "Voting weighted modified hausdorff distance through multiscale space for automatic image-map registration, " in *Proc. 18th International Conference on Pattern Recognition*, Vol. 4,pp. 837–840, 2006.

[8] L. Tian and S. Kamata, "Diffusion geodesic path: A common feature for automatic image-map registration, " in *Proc. Symposium on Signal Processing and Information Technology*, pp. 944–949, 2006.

[9] C. E. Shannon, "A mathematical theory of communications," *Bell System Technical Journal*, Vol. 27,pp. 379–423 and 623–656, 1948.

[10] H. Sagan, "A three-dimensional Hilbert-curve," *Journal on Math Ed. Sc. Tech.*, Vol. 24,pp. 541–545, 1993.

[11] J. Zhang and S. Kamata, "A Pseudo-Hilbert Scan for Arbitrarily-Sized Cuboid Region," in *Proc. Symposium on Signal Processing and Information Technology*, pp. 920–925, 2006.

[12] S. Kamata, M. Niimi, and E. Kawaguchi, "A method of an interactive analysis for multi-dimensional images using a Hilbert curve," *IEICE Trans.*, Vol.J77-D-II, No. 7, pp. 1255–1264, 1999.

[13] Y. Wang and H. Kuroda, "Hilbert scanning search algorithm for motion estimation," *IEEE Trans. Cir. and Sys. for Video Tech.*, Vol. 9, pp. 683–691, 1999.

[14] B. Moghaddam, K.J. Hintz, and C.V. Stewart, "Space filling curves for image compression," in *Proc. SPIE Conf. Automatic Object Recognition*, Vol. 1471, pp. 414–421, 1991.

[15] A.C. Ansari and A. Fineberg, "Image data ordering and compression using Peano scan and LOT," *IEEE Trans. Consumer electron*, Vol. 38, pp. 436–445, 1999.

[16] L. Tian, S. Kamata, K. Tsuneyoshi, and H.J. Tang, "A fast and accurate algorithm for matching images using Hilbert scanning distance with threshold elimination function," *IEICE Trans.*, Vol. E89-D, No. 1, pp. 290-297, 2006.

[17] L. Tian, S. Kamata, "Automatic Image-Map Alignment by Maximization of Edge-Based Code Mutual Information," in *Proc. of IEVC2007: IIEEJ Image Electronics and Visual Computing Workshop*, 2007.