Image Registration Based on Normalized Cross Correlation and Discrete Cosine Transform

PETAR RUTESIC
University of Belgrade
Faculty of Mathematics
Studentski trg 16, Belgrade
SERBIA
petar.rutesic@gmail.com

ZORANA STOSIC
University of Belgrade
Faculty of Mathematics
Studentski trg 16, Belgrade
SERBIA
stosiczorana@gmail.com

Abstract: There have been great advancements in recent years regarding computer vision, medical imaging, cartography, astronomy and similar image acquisition methods. Therefore, there is certainly need for efficient image registration techniques. Image registration is a process in which two images, that represent the same scene but are captured from different angles, sensors or time periods, are geometrically aligned to each other. This is one of the fundamental image processing techniques and is very useful when combining images captured from various sensors or time periods. The basic concept is to find points of interest by comparing unregistered image with source image. Points of interest are parts of source image that are highly similar to unregistered image. In this paper, points of interest are obtained by two approaches - by using normalized cross correlation (NCC) and discrete cosine transform (DCT). The proposed method was applied to satellite images. Tests have been successfully concluded even with high resolution images with some form of local distortion. Testing shows that DCT approach is quicker and more accurate.

Key–Words: normalized cross correlation, discrete cosine transform, image registration, root mean squared error, points of interest extraction

1 Introduction

Digital image processing represents one of the most widespread computer science area. In is used in medicine [1], [2], astronomy [3], biology [4], etc. Image registration is one of the common tasks in various applications. It is a process in which two images, that represent the same scene but are captured from different angles, sensors or time periods, are geometrically aligned to each other [5]. Main idea is to transform different datasets into one coordinate system. Image registration is one of the fundamental image processing techniques and is very useful when combining information from various sensors or when detecting differences in images obtained in different time intervals. In this paper, images are classified as source images and unregistered images. Source images were already acquired, and it is known when and how they were acquired what is shown on them while unregistered images need to be located in source images.

Main idea is to align source and unregistered images, by comparing unregistered images with areas of source image and extracting areas with high similarity [6]. Image registration process consists of a couple of basic steps [7]. Firstly, features extraction which is a process of identification of similar image parts according to some metric. Image feature extraction is used in various applications such as computer aided diagnostics [8], forgery detection [9], etc. It can be dot, line, angle, square etc. After this step, it is suppose to match corresponding attributes, i.e. match features of one images with features of another one. Next step is detecting geometrical transformations required so that source and unregistered images can be aligned to each other (rotation, translation, scaling etc.). Last step is image aligning which align both source and unregistered image to each other.

This paper focuses on features detection by using two approaches based on normalized cross correlation and discrete cosine transform (DCT) as regions features. DCT is used for different applications with image processing such as retina blood vessels detection [10], JPEG compression [11] and other. In [7], image registration method based on DCT coefficients was proposed. In this paper we added several steps that were customized. Unregistered image is compared to nonoverlapping blocks of source image, but unlike in [7], overlapping blocks are also considered. When using DCT, different sets of DCT coefficients are considered, while in [7] all DCT coefficients are used.

2 The proposed algorithm

In both approaches, input data for algorithm is the same: sample (part or whole unregistered image) with size $N \times M$ and source image with size $A \times B$. However, image blocks from source images are extracted differently. When using NCC or DCT algorithm by default splits source image into nonoverlapping blocks with the size of a sample, $N \times M$. Specially, when using Normalized Cross Correlation, algorithm can start from first pixel (point 0,0 in source image) and calculate similarity between source image and a sample for all blocks starting from all pixels which are in between $(0 : A-N)$ and $(0 : BM)$. This way, source image is split into overlapping blocks of $N \times M$ size.

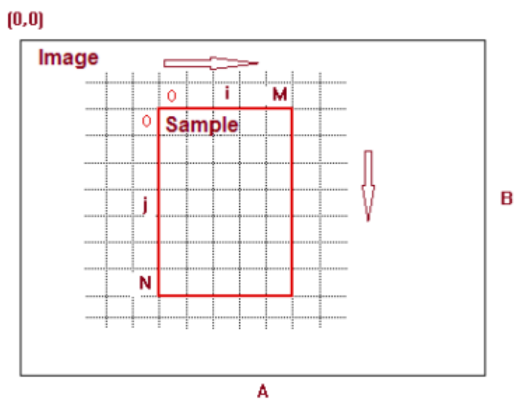


Figure 1: Overlapping blocks

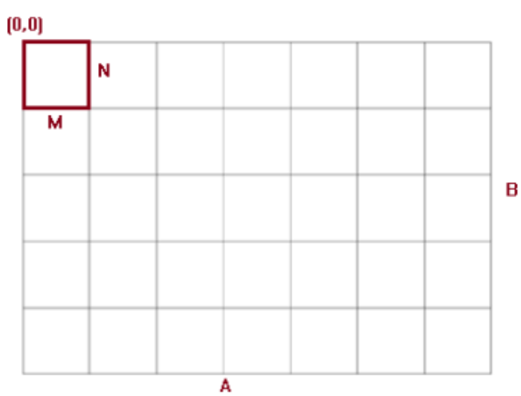


Figure 2: Non-overlapping blocks

Both of these methods are illustrated in Fig. 1 and Fig. 2. When image is split into non-overlapping blocks similarity is calculated between red and all gray blocks. When splitting image into overlapping blocks, similarity is calculated for blocks starting from every pixel (i, j) of source image, while i and j are less than $A - N$ and $B - M$. That upper limit is required because there would be problems if overlapping blocks are of size less than $N \times M$. This is obviously more demanding both from CPU usage and memory usage standpoint. CPU is utilized a lot more because Normalized Cross Correlation is calculated for every overlapping block, so the number of calculations is $(A - N)x(B - M)$, which is much more than when using non-overlapping blocks $(A/N) \times (B/M)$. This can be especially problematic if source images are high resolution one, and sample is small. Usage of overlapping blocks is also memory demanding because size of correlation matrix is (AN)x(B-M) instead of $(A/N) \times (B/M)$. However, this way sample image can be positively detected inside source image with almost 100% chance.

When using NCC, most similar block is the one which has greatest correlation with the sample. When using DCT, most similar block is the one that has lowest root-mean-square error (RMSE) with the sample. At the end of algorithm, most similar block from source image is extracted.

3 NCC based approach

Area based methods concentrate more attention to feature matching than to their detection [7]. Therefore, this approach does not use feature detection step so the first step in image registration process (mentioned in introduction) is omitted. Methods like this are sometimes called co-relational and they combine both first and second step in image registration process, and that is feature detection and feature matching.

For calculating normalized cross correlation Eq. 1 is used. Indexes y and x represent sample and block of source image, respectively.

$$r = r_{yx} = \frac{C(y,x)}{\sqrt{V(y)V(x)}}, \tag{1}$$

where:

$$C(y,x) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x}), \tag{2}$$

$$V(x) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{3}$$

Figure 3 is algorithm flowchart when using NCC. First step is loading unregistered (sample) image. After that, user chooses if he wants to compare sample with overlapping or non overlapping blocks of source image.

After NCC is calculated for selected sample and blocks from source image, block with the best correlation score is selected. Source image now has to be somehow transformed so that the matching block is more noticeable inside source image e.g. in implementation of this algorithm, source image is made

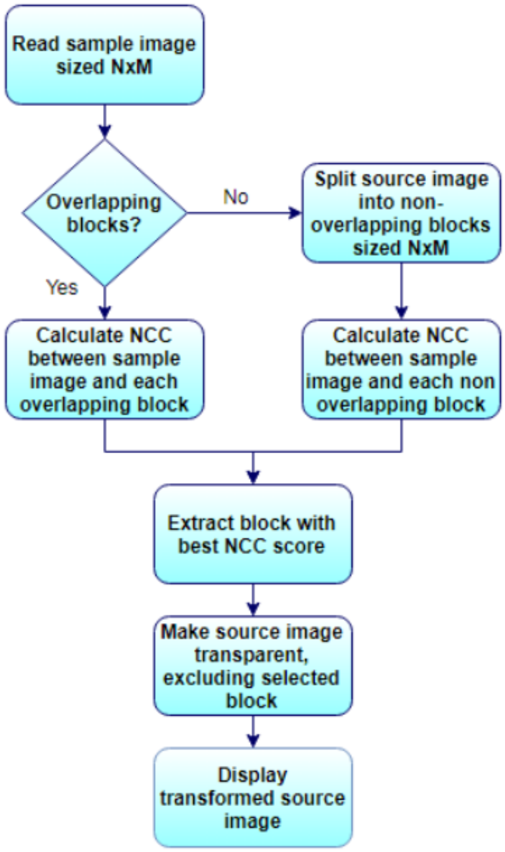


Figure 3: NCC based algorithm flowchart

more transparent, while chosen image block stays the same. That way, best block from source image is more accentuated. At the end, transformed source image is shown.

4 DCT based approach

Discrete cosine transform is highly related to discrete Fourier transform. This is a separable linear transform, which means that the two-dimensional transform is equivalent with one-dimensional DCT calculated on single dimension followed by a one-dimensional DCT in the other dimension. To calculate DCT coefficients the following equation is used:

$$D(u,v) = (\frac{2}{N})^{\frac{1}{2}}(\frac{2}{M})^{\frac{1}{2}}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} C(i)C(j)I(i,j) \quad (4)$$

$$\cos[\frac{\pi * u}{2 * N} * (2i+1)] \cos[\frac{\pi * v}{2 * M} * (2i+1)],$$

where N and M is the size of image, $I(i,j)$ is value of (i,j) pixel and

$$C(i) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } i = 0 \\ 1 & otherwise. \end{cases} \quad (5)$$

The difference between Discrete Fourier Transform and Discrete Cosine Transform is that DFT is calculated using sines and cosines while DCT uses only cosines. Symbol f(i,j) marks intensity of pixel at location (i,j) while F(u,v) is a value of DCT coefficient in DCT matrix.

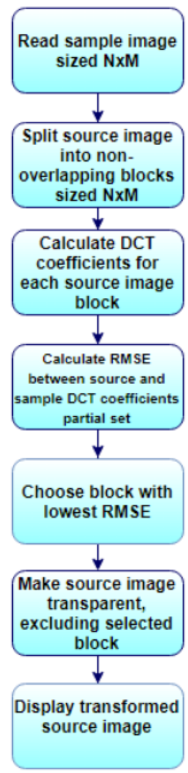


Figure 4: DCT based algorithm flowchart

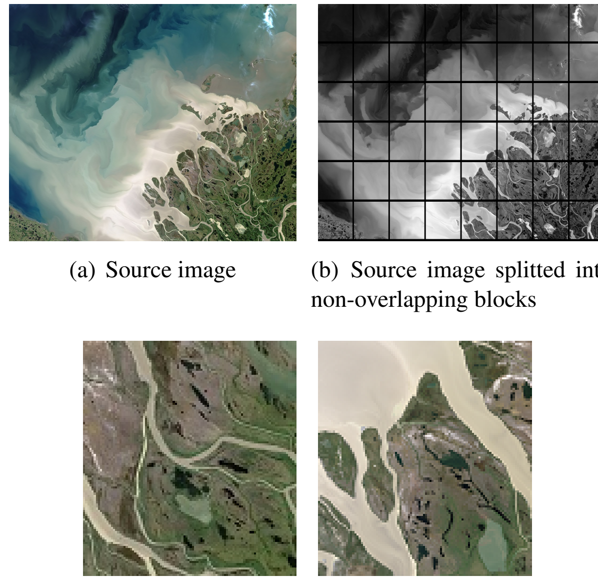
Main flowchart of the proposed DCT based method for image registration is given in Fig. 4. First step is to load source image and sample (unregistered image). After that, source image is split on non overlapping blocks, each sized $N \times M$, and DCT is calculated for each one of them. In implementation of this method, DCT is calculated by the following equation:

$$D = TMT' \quad (6)$$

where T is matrix of dimensions 8x8, T' is transposed matrix T. Size of matrix M is also 8x8, and it is one part of entire NxM block of source image. Each element of matrix M contains values of red channel from pixels of corresponding block, so that only one channel or grayscale image is considered.

After DCT is calculated, similarity measure has to be used so that blocks from source image similar to sample image can be detected. That measure is root-mean-square error. RMSE is calculated between sample image and each block of source image. In implementation, only low and middle frequency elements from DCT matrix have been used for RMSE calculation, because they carry most of image data. Block with lowest RMSE is chosen.

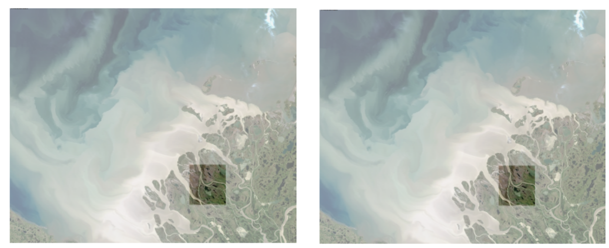
Source image now has to be somehow transformed so that the matching block is more noticeable inside source image e.g. in implementation of this algorithm, source image is made more transparent, while chosen image block stays the same. That way, best block from source image is more accentuated. At the end, transformed source image is shown.



(a) Source image (b) Source image splitted into non-overlapping blocks

(c) Sample 1 to be registred (d) Sample 2 to be registred

Figure 5: Test example image



(a) NCC based aproach (b) DCT based aproach

Figure 6: Image registration by our proposed method

5 Results

The proposed method was implemented in Java programming language. Public domain satellite images were used for testing, and they can be obtained at USGS/NASA Landsat web image gallery [12]. Images from various sizes were chosen. For same satellite images, tests with NCC and DCT have been conducted. Tests were running on a PC with Intel's i3 1.9 Ghz CPU and 4 GB of RAM.

DCT based approach seems to be a lot faster than NCC one, but results obtained with usage of non-overlapping blocks of source image and NCC are much more precise. Time spent for calculating RMSE is reduced by using only lower frequency coefficients of DCT matrix. However, false positives (wrong block being detected) can occur when not all DCT coefficients are used. Further examples are the results obtained.

6 Conclusion

Image registration has very important role when aligning two or more images that represent same scene captured in different time periods or from different angles is considered. Time distance can be measured by days, months, years etc.

Image algorithms that use Normalized Cross Correlation (NCC) or Discrete Cosine Transform (DCT) have been proposed. In NCC based approach, algorithm splits source image on overlapping or non-overlapping blocks of same size as sample (unregistered) image, NxM. This image registration method is used for satellite images. In DCT based approach, algorithm splits image into non-overlapping blocks and uses only lower frequency DCT coefficients.

After comparison of algorithms, DCT with usage of only lower frequency components proved to be fastest. However, false positives can occur on some of the satellite images used for testing.

Algorithm that uses NCC with overlapping blocks is the slowest one, but it is almost 100% accurate. However, execution time becomes problem when source image is high resolution and sample is small.

NCC based algorithm that uses non-overlapping blocks is faster than NCC with overlapping, but its not 100% accurate. Results vary and for some images false positives may occur.

Acknowledgements: This research is supported by the Ministry of Education, Science and Technological Development of Republic of Serbia, Grant No. III-44006.

References:

[1] E. Tuba, M. Tuba, and R. Jovanovic, "An algorithm for automated segmentation for bleeding detection in endoscopic images," in *International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 4579–4586.

[2] A. Stojak, E. Tuba, and M. Tuba, "Framework for abnormality detection in magnetic resonance brain images," in *24th Telecommunications Forum TELFOR*. IEEE, 2016, pp. 687–690.

[3] N. Grice, C. Christian, A. Nota, and P. Greenfield, "3D printing technology: A unique way of making hubble space telescope images accessible to non-visual learners." *Journal of Blindness Innovation & Research*, vol. 5, no. 1, 2015.

[4] B. Cook, D. Proctor, R. Bromberg, N. E. LaPointe, S. C. Feinstein, and L. Wilson, "Digital quantification of neurite outgrowth and retraction by phase-contrast microscopy: A tau perspective," in *Methods in cell biology*. Elsevier, 2017, vol. 141, pp. 217–228.

[5] E. Tuba, M. Tuba, and E. Dolicanin, "Adjusted fireworks algorithm applied to retinal image registration," *Studies in Informatics and Control*, vol. 26, no. 1, pp. 33–42, 2017.

[6] M.-S. Lee, M. Shen, A. Yoneyama, and C.-C. J. Kuo, "DCT-domain image registration techniques for compressed video," in *IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2005, pp. 4562–4565.

[7] R. B. Karani and T. K. Sarode, "Image registration using discrete cosine transform and normalized cross correlation," in *Proceedings on International Conference and workshop on Emerging Trends in Technology*, 2012, pp. 28–34.

[8] E. Tuba, I. Ribic, R. Capor-Hrosik, and M. Tuba, "Support vector machine optimized by elephant herding algorithm for erythemato-squamous diseases detection," *Procedia Computer Science*, vol. 122, pp. 916–923, 2017.

[9] I. Tuba, E. Tuba, and M. Beko, "Digital image forgery detection based on shadow texture features," in *24th Telecommunications Forum (TELFOR)*. IEEE, 2016, pp. 679–682.

[10] E. Tuba, L. Mrkela, and M. Tuba, "Retinal blood vessel segmentation by support vector machine classification," in *27th International Conference Radioelektronika*. IEEE, 2017, pp. 1–6.

[11] E. Tuba, M. Tuba, D. Simian, and R. Jovanovic, "JPEG quantization table optimization by guided fireworks algorithm," in *International Workshop on Combinatorial Image Analysis, LNCS*, 2017, vol. 10256, pp. 294–307.

[12] NASA, "Landsat image gallery." [Online]. Available: https://landsat.visibleearth.nasa.gov/