

NSCT edge Enhancement for SIFT key points extraction

Abdelkrim Ghaz¹, Kidiyo Kpalma² and Abdennacer Bounoua¹

¹Department of electronic, Faculty of Technology, Djillali Liabes University, Sidi Bel Abbès, Algeria

²INSA de Rennes, IETR UMR-6164, Université Européenne de Bretagne, France

Abstract: Image registration is a key step for matching or mosaicing two or more images taken at different times, and/or with different sensors, hence the need for automatic methods arises.

In this work, we present an efficient registration method based on the Non-Subsampled Contourlet Transform (NSCT) combined with the Scale Invariant Feature transform (SIFT) to extract robust local control points. Because NSCT is a shift-invariant multidirectional transform, it is used to extract edges at both spatial and directional resolutions.

A comparative study is established between SIFT and NSCT-SIFT and experimental results show clearly that the proposed method (NSCT-SIFT) improves the registration accuracy.

Keywords: Registration; SIFT; NSCT, key points extraction, edge enhancement

I. INTRODUCTION

Image Registration is the process of estimating an optimal transformation between two images [1]. It is an important task in many fields such as remote sensing [2] and medical imaging in particular image mosaicking [3]. In recent years, many image registration techniques have been developed based essentially on feature points detection, such as the SIFT method [4]. However, this technique is not very efficient when searching for identical point-pairs and eliminating mismatching. Furthermore, this method suffers from shifting that results from down sampling in the space scale extrema detection stage; in contrast NSCT [5] is fully shift-invariant and robust to noise. Based on the aforementioned issues, we propose a new approach combining SIFT and NSCT. In the proposed method we first use NSCT to extract the salient edges of images because this transform captures significant image features across its multi-directional and multi-resolution decomposition; then we add the edge map to the original image and apply the SIFT technique on the output.

II. OVERVIEW OF NSCT

The contourlet transform offers a multiresolution and multi-directional decomposition; it is introduced by Do and Vetterli [6]. Compared with the traditional wavelet, contourlet transform represents better the edges and other singularities along curves. However the contourlet transform has the propriety of being shift variant which is a handicap for some image processing applications. Cunha et al [7] proposed NSCT which is a fully shift-invariant version of the contourlet transform, and a multidirectional expansion that has a fast implementation. It is based on a nonsubsampled laplacian pyramid structure flowed by nonsubsampled directional filter banks. Figure 1 indicates the basic block diagram of the NSCT transform

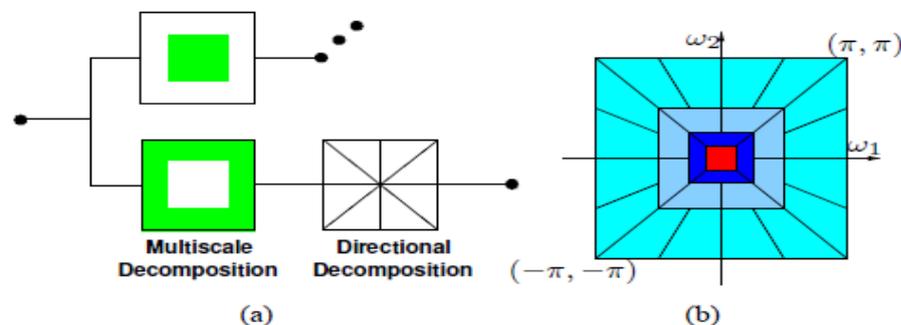


Fig. 1 The nonsubsampled contourlet transform

(a) Block diagram. First, a nonsubsampled pyramid splits the input into a lowpass subband and a highpass subband. Then a nonsubsampled DFB (directional filter banks) decomposes the highpass subband into several directional subbands. The scheme is iterated repeatedly on the lowpass subband.

(b) Resulting frequency division, where the number of directions increases with the frequency

III. OVERVIEW OF SIFT

The Scale Invariant Feature Transform (SIFT) is an image descriptor for image-based matching developed by David Lowe (1999, 2004). This descriptor as well as related image descriptors is used for a large number of purposes in computer vision related to point matching between different views of a 2-D scene. Experimentally, the SIFT descriptor has been proven to be very useful in practice for image matching and object recognition [4]

The following computation steps are used to generate the set of image features:

A. Detection of scale-space extrema

The first step of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian (DoG) function to identify potential key points that are invariant to scale and orientation. The DoG detector detects points in the Laplacian scale space. The scale space of an image is defined as a function, $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y, \sigma) \quad (1)$$

Scale-space extrema, which are extracted from the difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$, is defined as

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2)$$

This could be computed from the difference of two nearby scales separated by a constant multiplicative factor k . Then comparing against the points in its neighborhood, the key point is compared against its 26 neighbors in $3 \times 3 \times 3$ cube. If the pixel is a local maximum or minimum, it is selected as a candidate keypoint.

B. Accurate Keypoint Localization.

The keypoints localization includes point refinement by their instability in order to achieve their location accuracy. When the initial keypoints are found, there are a lot of methods to eliminate points that are useless for characterizing the image content. These are points that have a low contrast and are poorly localized along an edge. To eliminate the low contrast keypoints some threshold is chosen and each keypoint from the initial set the value of which is lower than this threshold is eliminated.

The difference-of-Gaussian function will have a strong response along edges. The eigenvalues of Hessian matrix are proportional to the principal curvatures and so the Hessian matrix of Difference of Gaussian scale-space can be used.

$$H = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} \quad (3)$$

If α is the eigenvalue with the largest magnitude and β be the smaller one. Then, the sum of the eigenvalues can be computed from the trace of H and their product from the determinant:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (4)$$

$$Det(H) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta \quad (5)$$

Thereby, we only need to check:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (6)$$

David Lowe assume $r=10$, that's to say, the experiments eliminate keypoints that have a ratio between the principal curvatures greater than 10.

C. Orientation assignment

In order to achieve invariance to the image rotation, each keypoint has to be assigned orientation based on the processed image visual properties. When image descriptor carries keypoints together with their respective orientations, then any possible rotations of the image don't affect the image comparison process. The keypoint

orientations are computed by using such image pixel features like gradient magnitude and orientation that are obtained by computing the pixel differences.

Pixel gradient magnitude and orientation are given as follows:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (7)$$

And

$$\theta(x, y) = \tan^{-1}((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1))) \quad (8)$$

An orientation histogram is formed. Peaks in the orientation histogram correspond to dominant directions of local gradients

D. The local image descriptor

The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination. According to the above operation, image location, scale, and orientation are assigned for each keypoint. A keypoint descriptor is created by computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location. The experiments use a $4 \times 4 \times 8 = 128$ element feature vector for each keypoint. Finally, for the sake of reducing the effects of illumination change, the vector is normalized.

IV. PROPOSED METHOD

In our improved algorithm of image registration, we use combination of NSCT to enhance [8] edges and SIFT Transform for more and accuracy control points extraction, other methods combine SIFT and CONTOURLET Transform[9] to choose the fine matching points to realize the image registration

The proposed key points extraction is summarized in the following steps:

1. Apply the NSCT decomposition and we consider 3 decomposition levels and 4 directions.
2. Compute the maximum between the 4 subbands using equation (9), this operation is called “maxima edge”.

$$S_{ij} = \arg \max_{\rho \in [B_1^n(i, j), B_2^n(i, j), B_3^n(i, j), B_4^n(i, j)]} |\rho| \quad (9)$$

With B_k represents different subband directions $k = 1, 2, 3, 4$

n Indicates the scale level.

3. At each level, the output of “maxima edge” is weighted by respectively α, β, γ and the original image by Θ , such that the condition $\alpha + \beta + \gamma + \Theta = 1$ is respected.
4. The result of the edge map is added to the original image to enhance image edges.
5. Subsequently SIFT is applied to the edge-enhanced image to extract the keypoints.

Figure 2 illustrates a block diagram of the proposed combined NSCT-SIFT algorithm with 3 decomposition levels and 4 directions.

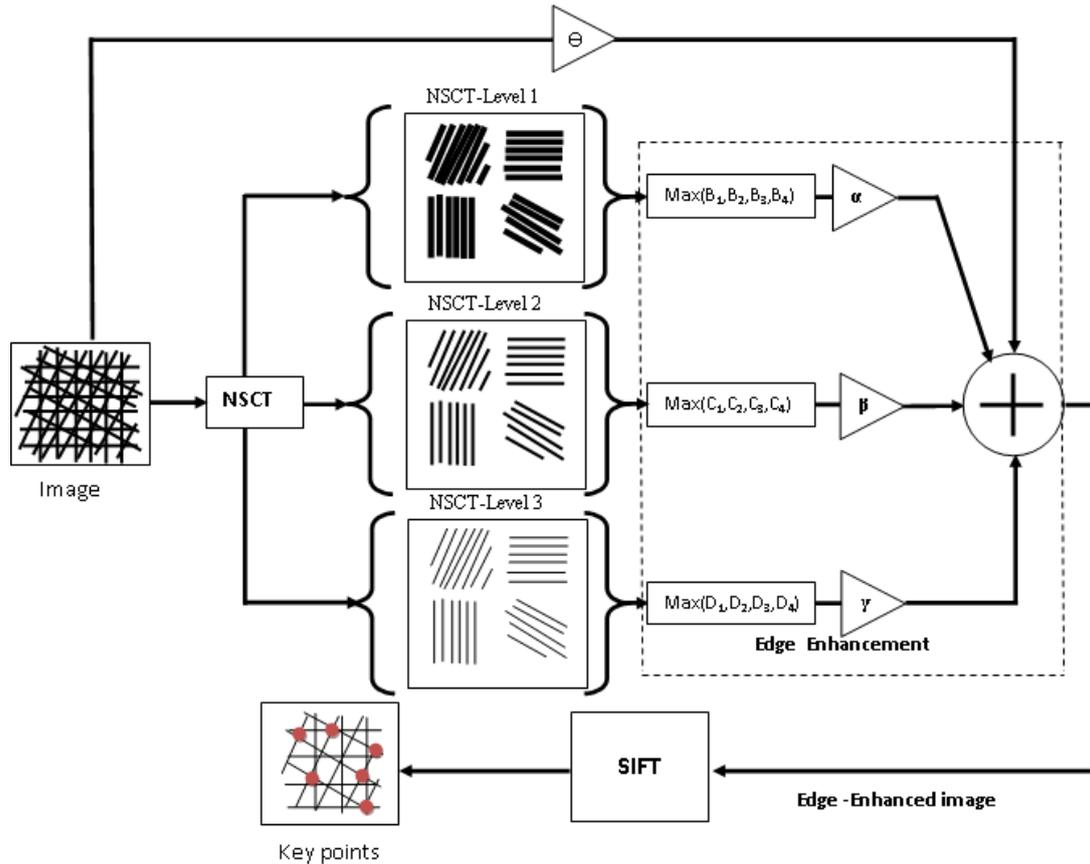


Fig. 2 block diagram of the proposed NSCT-SIFT algorithm to extract keypoints

V. EXPERIMENTAL RESULTS

The proposed algorithm is first evaluated by applying it to LENA image. Figure 3 shows the resulted keypoints obtained from SIFT applied to ensharped image.

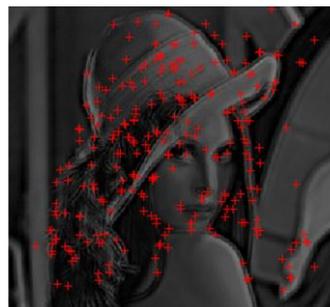


Fig. 3 Extracted Keypoints surimposed on the enhanced image

The experimental tests were performed using Lena and boat images . In these first experiments, the transformed or input images to be corrected are simply the reference images rotated by a 7 degree angle of rotation. The sensed images are resampled using the bilinear interpolation. Figure 4 illustrates the registered images obtained using the NSCT-SIFT and SIFT methods. The pair of images to be compared is shown with a size of (256 × 256) pixels.

It is to note that the NSCT-SIFT method gives a better PSNR than that of the SIFT one after registration.

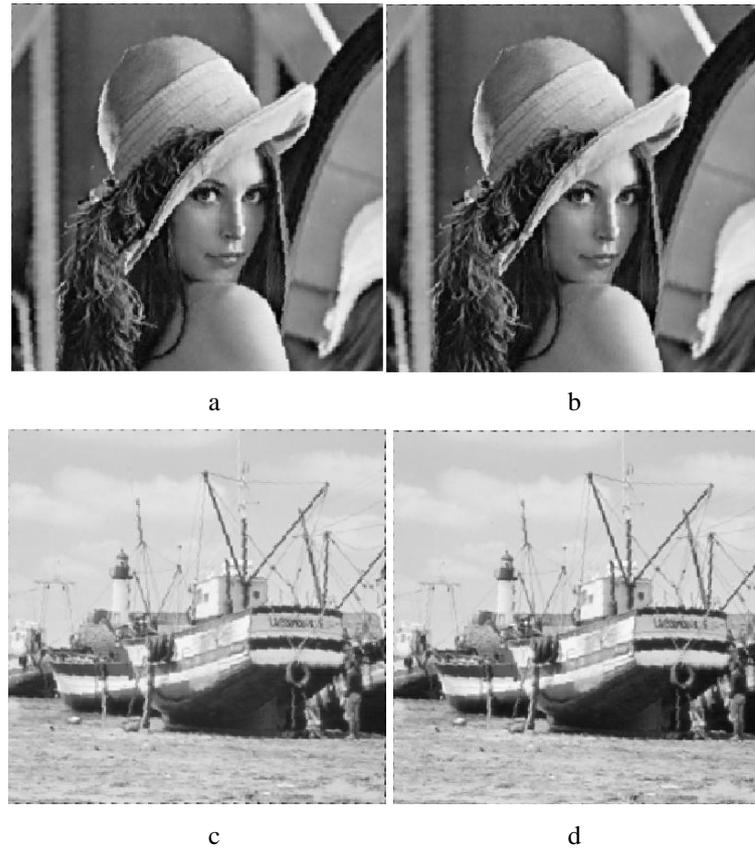


Fig. 4. Registered LENA and BOAT images (a) using the SIFT method, (PSNR = 34.1389) (b) using the NSCT-SIFT method, (PSNR = 38.5915) (c) using the SIFT method, (PSNR = 41.3492) (d) using the NSCT-SIFT method, (PSNR = 42.8109)

Some metric parameters are used to evaluate the quality of registered images namely the correlation coefficient CC between two images as defined in Equation (10), and the root mean square error (RMSE) as defined in equation (11)

$$CC = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} ((X_{ij} - m_x) \times (Y_{ij} - m_y))}{\sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (X_{ij} - m_x)^2 \times \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (Y_{ij} - m_y)^2}} \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N \times N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (X_{ij} - Y_{ij})^2} \quad (11)$$

Where X and Y are images of size $N \times N$, while X_{ij} and Y_{ij} denote the pixel values in X and Y respectively.

To illustrate the robustness and accuracy of our registration method, we applied a rotation of 10 degrees to the Reference Image with Gaussian noise [10] with mean 0 and variance 0.02. Experimental results show that the points calculated by the method NSCT sift gives better registration than those obtained by the sift transform where you can see a duplication of some structures after superposition of the original image with the registered image as demonstrated by figure 5.

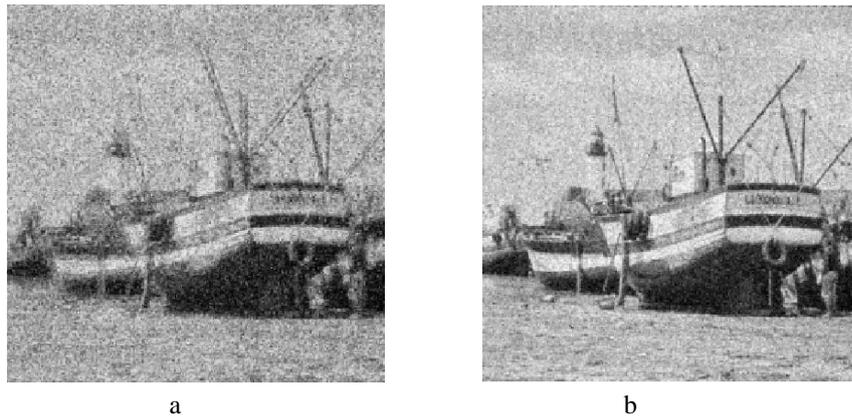


Fig 5: Registered image overlaid on Reference image **a)** with SIFT method **b)** with NSCT-SIFT method

Then, our method has been tested using various values of noise variances. For each variance value, the experiments are conducted ten times; and each time the PSNR is calculated. After that, for each variance value, the mean of the ten PSNR values is saved. Figure 6 represents PSNR values obtained for different variance values using the SIFT and NSCT-SIFT (our method) methods. From this figure, it is obvious that the proposed algorithm is more accurate and efficient compared with the SIFT algorithm. The good PSNR values are always obtained for the proposed NSCT-SIFT algorithm.

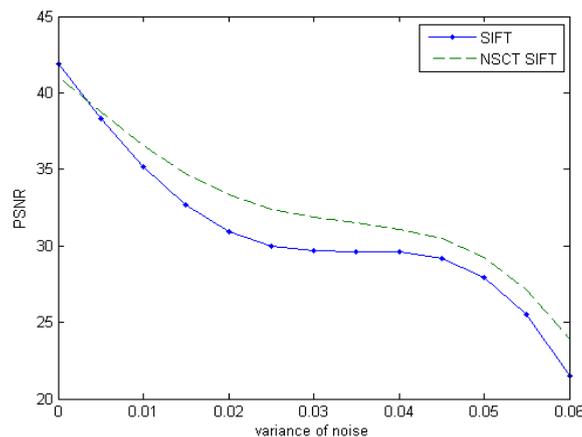


Fig 6: comparison of PSNR after registration between SIFT and NSCT-SIFT transforms

To test our proposed technique, we have used different types of satellite images such as SPOT and IKONOS as well as radar images SAR. The pair of images to be compared is shown with a size of (256×256) and are simply be rotated by a 7 degree angle.

It is also to note that the NSCT-SIFT method gives a better PSNR than that of the SIFT one after registration, as illustrated in figure 7.

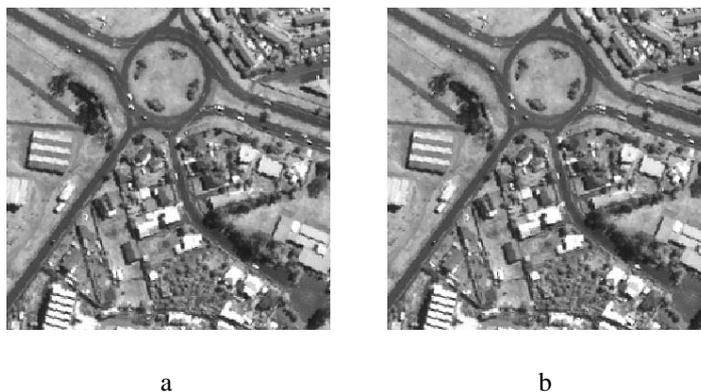


Fig.7: Registered IKONOS image using **a)** the SIFT method, PSNR =35.1609 **b)** the NSCT-SIFT method, PSNR =38.5511

The different analytical results of the registration methods on all image sets are recapitulated in Table 1. Two measures are considered to determine the accuracy of registration, the correlation coefficient and the root mean square error. The results are compared with those obtained in [11].

Table 1: Analytical results for different registration methods

		Techniques				
		Wavelet	NSCT1	NSCT2	SIFT	NSCT-SIFT
SPOT	Corr.rati	0.8298	0.9863	0.9840	0.9908	0.9914
	RMSE	0.0462	0.0137	0.0147	0.0111	0.0108
IKONOS	Corr.ratio	0.6935	0.9562	0.9468	0.9627	0.9731
	RMSE	0.1245	0.0504	0.0549	0.0461	0.0393
SAR	Corr.ratio	0.5622	0.9474	0.9294	0.9517	0.9569
	RMSE	0.1549	0.0608	0.0703	0.0520	0.0491

Table 1 shows clearly that both proposed methods SIFT and NSCT-SIFT perform better in terms of correlation and RMSE than the wavelet, NSCT1 and NSCT2 methods proposed in [11]. In the NSCT1 method, authors decompose the two images to be compared into several levels with different directions. At each level of the decomposition, the correlation ratio between corresponding sub-band images of the reference and input images is successively computed and maximized using GAs. The GA's process is applied on all the obtained bands. However, in the NSCT2 approach, the GA's process is performed on only the approximation of the original image sub-band (LL).

VI Conclusion

This paper has presented a powerful method for feature points detection and extraction in spatial images. This method is based on the combination of two well known techniques NSCT and SIFT. This transform is able to capture significant image features across multi-directional and it's more stable for geometric transformation.

Experimental results show that using SIFT with NSCT decomposition at the third scale level and 4 directions gives the best results. The correlation coefficient and RMSE illustrate minimum registration error for NSCT-SIFT which is suitable in remote sensing image registration even in the presence of noise.

References

Journal Papers:

- [1]. Zitová, B.; Flusser, I. "Image registration methods" A survey. *Image Vision Comput.* 2003, 21, p 973-1000.
- [2]. Bentoutou, Y.; Taleb, N.; Kpalma, K.; Ronsin, J. An automatic image registration for applications in remote sensing. *IEEE Trans. Geosci. Remote Sens.* 2005, 43, p 2127-2137.
- [3]. Do, M.N.; Vetterli, M. "Contourlet" in *Beyond Wavelets*, Academic Press, New York 2003.
- [4]. Cunha et Al. The nonsampled contourlet transform: theory design and applications. *IEEE Transactions on image Processing*, 15(10), 2006, p.3089-3101.
- [5]. Jignesh Sarvaiya, Suprava Patnaik & Hemant Goklani "Image Registration using NSCT and Invariant Moment" *International Journal of Image Processing (IJIP)*, Volume (4): Issue (2) , May 2010, p.119-130.
- [6]. Meskine, F.; Chikr el mezouar, M. and Taleb, N. "A Rigid Image Registration Based on the Nonsampled Contourlet Transform and Genetic Algorithm", *SENSORS*, 10(9):. Sept.2010..p 8553-8571.

Proceedings Papers:

- [7]. Zhao Xiuying ; Wang Hongyu ; Wang Yongxue " Medical Image Seamlessly Stitching by SIFT and GIST" *E-Product E-Service and E-Entertainment (ICEEE)*, International Conference on , 2010, p 1- 4,
- [8]. David, L. "Object recognition from local scale-invariant features ", *Proceedings of the International Conference on Computer Vision*, vol. 2, 1999, p. 1150-1157.
- [9]. Zhou, J.; Cunha, A.L. and Do, M.N., "The Nonsampled Contourlet Transform: Construction and Application in Enhancement". In *Proceedings of IEEE Int. Conf. on Image Processing, ICIP 2005*, (1) , 2005 p. 469-472.
- [10]. Feng Zhou ; Xiuli Ma ; Yang Li ; Xiaojun Zhou "Medical image enhancement based on NSCT" *IET International Conference on, (ICSSC 2013)* .p 166 - 169
- [11]. Pengrui Qiu and Ying Liang "The Improved Algorithm of Remote Sensing Image Registration Based on SIFT and CONTOURLET Transform " *Software Engineering and Service Science (ICSESS)*, 4th IEEE International Conference on 23-25 May 2013 p 906 - 909