

Scientific Journal of Impact Factor (SJIF): 3.134

International Journal of Advance Engineering and Research Development

Volume 2, Issue 5, May -2015

Infrared-Visible Video Registration: Survey

Pavan Modi¹, Dr. Darshak G. Thakore², Dr. Tanmay D. Pawar³

¹Student (PG), Computer Engineering, Birla Vishvakarma Mahavidyalaya, Vallabh Vidyanagar, ²Head Of Department, Computer Engineering, Birla Vishvakarma Mahavidyalaya, Vallabh Vidyanagar, ³Head Of Department, Electronics Engineering, Birla Vishvakarma Mahavidyalaya, Vallabh Vidyanagar,

Abstract-Image Registration is method is proposed to perform image registration in visible and infrared pair of video Sequences for multiple targets. This has led to the development of variety of applications ranging from medical imaging, computer vision, remote sensing, and astrophotography. There are five steps for video registration: Video converted into frames, feature detection and matching, transform of visible and infrared frames based on matching points, registered frames, generate videos. Speeded Up Robust Feature (SURF), SIFT, PCA-SIFT, Normalized cross correlation etc. are used for feature detection and matching. Based on matching points, images are transferred using affine transform, homography transformed etc.

Keywords-Image Registration, visible video, infrared video, SURF

I. INTRODUCTION

Image Registration is a process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors.

The present variations between images are introduced due to different imaging conditions. Image registration is may be a crucial step all told in image analysis tasks within which the ultimate info is gained from the mixture assorted information sources like in image fusion, amendment detection, etc. toextract more information about an object of interest in an image. Visible camera provides information about the visual context of the objects in the scene, but under poor light conditions only limited information is captured. On the other hand infrared provides enhanced contrast and rich information about the object when there is less light, especially in the darkenvironment [1]. Example of such different capturing information is shown in Figure 1.

Infrared-visible image registration is a very challenging problem since the thermal and visible sensors capture different information about a scene [1]. The infrared captures the heat signature emitted by objects, while the visible captured the light reflected by objects. Due to this difference, the correspondence between the visible and the infrared is hard to establish as local intensities or textures do not match, as can be seen in the Figure 1.



(a)



(b)

Fig. 1. Different image information captured of a same scene by (a) visible camera (b) and infrared camera, respectively[14].

There are various techniques used for feature detection and matching as Speeded up Robust Feature (SURF), Scale Invariant Feature Transformed (SIFT), Normalized cross correlation etc. SURF is faster than SIFT because it uses integral images and box filter to detect feature descriptor which are scale, rotation, and illumination invariant. Based on feature points, images are transformed using affine transformed, holography transformed.

II. FEATURE DETECTION AND MATCHING

D. G. Lowe(2004) presented that SIFT algorithm is a method to extract and describe feature points, which is robust to scale, rotation and change in illumination.

There are four steps to implement the SIFT algorithm [12]:

- 1. Scale-space extrema detection
- 2. Feature point localization
- 3. Orientation assignments
- 4. Feature point descriptor

Y. Ke and R. Sukthankar (2004) introduced PCA (Principle component analysis)-SIFT which accepts the same input as the standard SIFT descriptor [13]: the sub-pixel location, scale, and dominant orientations of the key point. We extract a 41×41 patch at the given scale, centered over the key point, and rotated to align its dominant orientation to a canonical direction. PCA-SIFT can be summarized in the following steps: (1) pre-compute an Eigen space to express the gradient images of local patches; (2) given a patch, compute its local image gradient; (3) project the gradient image vector using the Eigen space to derive a compact feature vector.

SURF: The search for discrete image point correspondences can be divided into three main steps.

- First, `interest points' such as corners, blobs, and T-junctions are selected at distinctive locations in the image.
- Next, using feature vector the neighborhood of every interest point is represented. This descriptor has to be distinctive and at the same time robust to noise, geometric and photometric deformations and detection displacements.
- Finally, between deferent images the descriptor vectors are matched based on a distance between the vectors, e.g. the Euclidean distance.

Integral Image or summed area tables is an intermediate representation of the image. It contains the sum of intensity values of all pixels in input image I within rectangular region formed by origin O=(0,0) and any point X=(x,y). It provides fast computation of box type convolution filters.

$$I_{\Sigma}(X) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i, j)$$
(1)

After computing integral image, only three operations (addition or subtraction) require for calculating sum of intensity values of pixels over any upright rectangular area.

The Hessian matrix $\mathcal{H}(X, \sigma)$ in X for a point X=(x,y) of an image I, at scale σ is defined as follows:

$$\mathcal{H}(\mathbf{X},\sigma) = \begin{bmatrix} \mathbf{L}_{xx}\left(\mathbf{X},\sigma\right) & \mathbf{L}_{xy}\left(\mathbf{X},\sigma\right) \\ \mathbf{L}_{xy}\left(\mathbf{X},\sigma\right) & \mathbf{L}_{yy}\left(\mathbf{X},\sigma\right) \end{bmatrix}$$
(2)

Where $L_{xx}(X,\sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2}g(\sigma)$ with the image I in point X and similarly for $L_{xy}(X,\sigma)$ and $L_{yy}(X,\sigma)$.

For scale-space analysis Gaussians are optimal, but in practice they have to be discretized and cropped. Under image rotations around odd multiples of $\frac{\pi}{4}$ this leads to a loss in repeatability. But in Hasian matrix due to the square shape of the filter repeatability is maximum around multiples of $\frac{\pi}{2}$.

In order to be invariant to image rotation, we identify a reproducible orientation for the interest points. Due to this, we first calculate the Haar wavelet responses in x and y direction within a circular neighborhood of radius 6s around the interest point, with scale s at which the interest point was detected. The sampling step s is scale dependent. In keeping with the rest, also the size of the wavelets are scale dependent and set to a side length of 4s. Therefore, we can use integral images for fast filtering again. To compute the response in x or y direction at any scale only six operations are needed.



Fig. 2: Haar wavelet filters to compute the responses in x (left) and y direction (right). The dark parts have the weight -1 and the light parts +1.^[7]

For the extraction of the descriptor, square region cantered around the interest point constructed and oriented along the orientation selected in the previous section. The size of this window is 20s. Examples of such square regions are illustrated in figure 3.



Fig. 3: Detail of the Graffiti scene showing the size of the oriented descriptor window at different scales.^[7]

The region is split up into smaller 4 x 4 square sub-regions regularly. This preserves important spatial information. For each sub-region, we compute Haar wavelet responses at 5 x 5 regularly spaced sample points. For reasons of simplicity, we call dx and dy, the Haar wavelet response in horizontal direction and dy the Haar wavelet response in vertical direction respectively (filter size 2s), see figure 4.



Fig. 4: To build the descriptor, an oriented quadratic grid with 4 x 4 square sub-regions is laid over the interest point (left). For each square, the wavelet responses are computed. The 2 x 2 sub-divisions of each square correspond to the actual fields of the descriptor. These are the sums dx, |dx|, dy and |dy| computed relatively to the orientation of the grid (right).^[7]

Then, the wavelet responses dx and dy are summed up over each sub-region and form a first set of entries in the feature vector. In order to bring in information about the polarity of the intensity changes, we also extract the sum of the absolute values of the responses, |dx| and |dy|. Hence, each sub-region has a four-dimensional descriptor vector V for its underlying intensity structure $V = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_x|)$. Concatenating this for all 4 x 4 sub-regions, these results in a descriptor vector of length 64. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

III. TRANSFORMATIONOF AN IMAGE

There are many methods are used for transformation of images as affine transformation, homography transformation etc. The affine transformation parameters can be calculated by coordinates of control points, and then geometrically transformation may be conducted for registered image.

[x′]	_ [cosī(a)	-sin(a)	0] []	$\left[\Delta x\right]$ (2)
[y']	- ^{s.} lsin@a)	cos (a)	$0] \begin{bmatrix} y \\ 1 \end{bmatrix}^{+}$	$+ [\Delta y] \dots (3)$

Another method is RANSAC is used for removing false matching pairs and transformation of image. The RANSAC calculation presented by Fischler and Bolles in 1981 is a vigorous estimator that is equipped for giving

great assessments from information defiled by expansive division of exceptions. It is an iterative and non-deterministic calculation which delivers a sensible outcome just with a certain likelihood, with likelihood expanding as higher cycles are permitted.

Algorithm 1 RANSAC:

1: Select randomly the minimum number of points required to determine the modelparameters.

2: Solve for the parameters of the model.

3: Determine how many points from the set of all points fit with a predefined tolerance \in .

4: If the fraction of the number of inliers over the total number points in the setexceeds a predefined threshold τ , reestimate the model parameters using all theidentified in liers and terminate.

5: Otherwise, repeat steps 1 through 4 (maximum of N times).

IV. CONCLUSION

Because the scenes are usually not planar, alignment needs to be performed continuously by extracting relevant common information of visible and infrared videos. SURF is robust and efficient approach for feature detection and matching because it is using integral image and box filter. RANSAC is better approach than affine transform for remove false matching pairs and transformed of image.

REFERENCES

- [1] Tanushri Chakravorty, Guillaume-Alexandre Bilodeau. "Automatic Image Registration in Infrared-Visible Videos", IEEE International Conference, March 2012.
- [2] Jun gong Han, Eric Pauwels, and Paul de Zeus." Visible and Infrared Image Registration Employing Line-Based Geometric Analysis", Springer-Verlag Berlin Heidelberg, pp. 114-125, 2012.
- [3] Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma, and Josep Ronsin," An Automatic Image Registration for Applications in Remote Sensing", VOL. 43, NO. 9, SEPTEMBER 2005.
- [4] Atousa Torabi, Guillau me Mass 'e, Guillau me Alexandre Bilodeau." An iterative integrated framework for thermalvisible image registration, sensor fusion, and people tracking for video surveillance applications" Canada, H3C 3A7.
- [5] Ms. Chandrakala.M, Mrs. R.Amsaveni. "Classification of Remote Sensing Image Areas Using Surf Features and Latent Dirichlet Allocation" Volume 3, Issue 9, September 2013.

- [6] Herbert Bay; Andress Ess, Tinne Tuytelaars,Luc Van Gool, "Speeded-Up robust (SURF)" Vol. 110, No. 3, pp. 346--359, June 2008.
- [7] Herbert Bay, Tinne Tuytelaars2, and Luc Van Gool121 ETH Zurich "SURF: Speeded Up Robust Features" Katholieke Universities Leuven.
- [8] Teodor Costachioiu, Rodica Constantinescu, Bashar AlZenk and Mihai Datcu, "Semantic Analysis Of Satellite Image Time Series", 20th European Signal Processing Conference (EUSIPCO 2012) Bucharest, Romania, August 27 - 31, 2012.
- [9] Darshana Mistry, Asim Banerjee," Review: Image Registration", International forum of researchers students and academeciants, Feb. 2012.
- [10] L. Zagorchev, A. Goshtasby, "A Comparative Study of Transformation functions for Non rigid Image Registration", "Transformation Functions for Image Registration", Image fusion systems research.
- [11] Utsav Shah, Darshana Mistry, Asim Banerjee, "Image Registration of Multi-View Satellite Images Using Best Feature Points Detection and Matching Methods from SURF, SIFT and PCA-SIFT", Journal of Emerging Technologies and Innovative Research, Vol. 1, Issue. 1, 2014
- [12] Konstantinos G. Derpanis, "Overview of RANSAC algoritham", May 2013.
- [13] L. Juan, O.Gwun(2009), "A comparison SIFT, PCA-SIFT and SURF", International Journal of Image Processing, Vol.-III, Issue(IV), pp. 143-152.
- [14] Videodatabase:- http://www.vcipl.okstate.edu/otcbvs/bench/Data/03/download.html