ON INTEREST POINT DETECTION UNDER A LANDMARK-BASED MEDICAL IMAGE REGISTRATION CONTEXT

*P. Martins and P. Carvalho**

CISUC, Department of Informatics Engineering, University of Coimbra
Pólo II, 3030-290 Coimbra, Portugal

ABSTRACT

A comparison and a performance evaluation of 3D interest point detectors based on the structure tensor matrix under a medical image registration context is presented. The study regards the distinctiveness and the repeatability rate exhibited by the detectors on medical images as the fundamental criteria for selecting a detector to be the support of a registration task. We empirically assess those requirements under the specificities of medical image analysis and registration.

Index Terms— Interest Points, Landmark-based Registration, Invariant Features, Medical Image Registration.

1. INTRODUCTION

Common solutions to computer vision problems involving tasks such as image registration, tracking, image retrieval or camera pose estimation often rely on low-level feature detection methods. The main motivation for that strategy has its basis on the intrisic properties of the features that let foresee an efficient and robust detection. Interest points, i.e., pixel locations exhibiting some singularity with respect to their neighborhoods have been largely studied in the literature and pointed out as appropriate features to be the support of image registration methods mainly due to their efficiency, distinctiveness and invariance (or co-variance) when dealing with geometric and photometric transformations [1]. In the context of medical image registration (MIR), interest point detection has been suggested as an alternative to the detection of anatomical landmarks (salient and accurately locatable points of the morphology of the visible anatomy [2]), avoiding, therefore, the need of user interaction. Regardless the modality, the dimensionality or the observed anatomy, medical images usually do not exhibit frequent considerably high variations in intensity as we often see in other types of images. Hence, one cannot expect interest points to be abundant in those images. However, the reduced number of such features reveals itself as an advantageous characteristic that should be taken into account when pondering the foundations of a MIR method; the distinctive locations irregularly present on images, assured by means of some criterion, provide a suitable underlying support for point-based registration. Generic comparisons of local features often refer to MIR as a task demanding an effective feature detection, however those studies do not consider its unique characteristics. The fundamental aim of this paper is to study 3D interest point detection in the MIR context, not by giving a comprehensive review of those algorithms under the specific environment, but by answering to questions that appear as essential on the choice of a particular detector: (i) how repeatable are interest points under certain conditions? (ii) How distinctive are interest points on medical images? (iii) Does MIR impose improved interest point detection methodologies? Thus, we set the structure tensor matrix of an intensity image, which is the basis of numerous 2D interest point detectors, as the starting point of our study and provide a mathematical framework to review interest point detection and to analyze it under the specificities of MIR. It is worth noting that the choice of the aforementioned solutions as a subject of analysis relies on their straightforward extension to any dimension $n \geq 3$ and on the well known high repeatability rate results in the presence of image rotation, translation, noise addition or illumination changes. The low computational complexity exhibited by those methods is another key feature to consider. That constraint is even more critical, for example, in computer guided surgery scenarios in which real-time automatic landmark detection is demanded.

For a better reading, Section 2 introduces some of the definitions, notations and remarks that will be followed throughout this document. Section 3 reviews four interest point detection functions based on the structure tensor matrix. In Section 4, we present and analyze the results of experiments required to assess the performance of the detectors. Finally, in Section 5, conclusions are driven.

2. PRELIMINARIES

Herein, the terms *interest point* and *landmark* will be adopted, the latter will refer to an interest point in a medical imaging context or some other reference point on a medical image. By considering the definition of interest point, several solutions can be regarded as interest point detectors, just as long they

*This research was supported by FCT (project PTDC/EEA-ACR/68887/2006), IIIUC and CISUC.

provide a reliable measure of the singularity of points with respect to their vicinities. Thus, a possible generic and coarse definition of an interest point detector could be the following:

Definition 2.1 ((Local) interest point detector) *Let $I : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be an image ($\mathcal{D}$ is a compact set). A (local) interest point detector is a real-valued function f defined over a given pixel location $\mathbf{x} \in \mathcal{D}$ and a certain window Ω about the aforementioned location, such that if $\mathbf{x}$ and $\mathbf{y}$ are interest points, it will be possible to define an equivalence relation by means of f such that the locations $\mathbf{x}$ and $\mathbf{y}$ belong to the same equivalence class.*

To dissipate the vagueness of definition 2.1, we can say that most of the detectors identify interest points as the locations where f is above a certain threshold. The structure tensor matrix - a representation of local image gradient information, - is known to be one effective description of local patterns.

Definition 2.2 (Structure tensor matrix) *Let I be a n-dimensional single channel image. The structure tensor matrix of I at the image point $\mathbf{x}$ is the $n \times n$ matrix T which results of the averaged outer product of the image gradients:*

$$T(\mathbf{x}) = (\omega * \nabla_{\mathbf{x}} \mathbf{I} \nabla_{\mathbf{x}}^{\mathbf{T}} \mathbf{I})(\mathbf{x}), \tag{1}$$

where ω is the weighting function and $\nabla_{\mathbf{x}} I$ is the gradient of I at $\mathbf{x}$.

The spectral structure of T captures the local signal changes along the principal directions: the order of magnitude of the eigenvalues is proportional to the gradient variation along them. Since the determinant of T, which is the product of all eigenvalues, exhibits large values at these locations, one can think of that operator as a detector that identifies interest points as the locations that are above a given threshold. In fact, most of the detectors based on the aforementioned matrix rely on the determinant or the trace of T.

3. INTEREST POINT DETECTORS

To simplify the notation, we neglect the dependence of the detectors on I and Ω. In the solutions described below, the gradient is obtained from a smoothed version of I: $L(\mathbf{x}, \delta) = (G_\sigma * I)(\mathbf{x})$, where G_σ denotes a Gaussian kernel with standard deviation σ. Moreover, we rewrite T as $A^T W A$, where $A = [a_{ij}] \in \mathbb{R}^{m \times n}$, with $a_{ij} = \frac{\partial}{\partial x_j} L(\mathbf{x_i}, \sigma_D)$; $\mathbf{x_i}$ are the points from the discretized support set of ω and $W = [w_{ij}] \in \mathbb{R}^{m \times m}$ denotes the diagonal weighting matrix, i.e., $w_{ii} = \omega(\mathbf{x} - \mathbf{x_i})$. In Eqs. (2-6), λ_i, $i = 1, \ldots, n$, represent the eigenvalues of T in non-decreasing order.

In [3], Noble suggests a modified version of the widely used Harris-Stephens detector [4] that does not contain the original tuning parameter, overcoming the need of manually tuning it:

$$f_1(\mathbf{x}) := \frac{\det(A^T W A)}{\epsilon + \operatorname{trace}(A^T W A)} = \frac{\prod_{i=1}^n \lambda_i}{\epsilon + \sum_{i=1}^n \lambda_i}, \tag{2}$$

where ϵ denotes an arbitrary small positive constant. By neglecting the trace of T, Rohr [5] proposes an alternative detector relying exclusively on the determinant of T :

$$f_2(\mathbf{x}) := \det(A^T W A) = \prod_{i=1}^n \lambda_i. \tag{3}$$

Both of the aforementioned measures imply the exclusion of points that convey a less discriminant information: those exhibiting a large difference between the eigenvalues. For that, the following threshold is used:

$$\frac{\det(A^T W A)}{(\frac{1}{n} \operatorname{trace}(A^T W A))^n} = \frac{\prod_{i=1}^n \lambda_i}{(\frac{1}{n} \sum_{i=1}^n \lambda_i)^n}. \tag{4}$$

The Shi-Tomasi algorithm [6] detects interest points via Eq. (5). By imposing $f_3(\mathbf{x}) \geq \lambda$, where $\lambda \in \mathbb{R}^+$ is a sufficiently large threshold, we are selecting points whose corresponding structure tensor matrix exhibits eigenvalues that do not differ by several orders of magnitude, which in terms of visual patterns, corresponds to select those that show high intensity variations in several directions.

$$f_3(\mathbf{x}) := \lambda_{min}(A^T W A) = \lambda_1 \tag{5}$$

The Kenney *et al.* detector [7] for the Schatten p-norm, with $p \in [1, \infty]$, identifies interest points using the following function:

$$f_4(\mathbf{x}) := \frac{1}{\|(A^T W A)^{-1}\|_p} = \frac{1}{\sqrt[p]{\sum_{i=1}^n \frac{1}{\lambda_i^p}}}. \tag{6}$$

The latter detector arises as an explicit attempt to select locations that convey a better repeatability in the presence of image rotations and translations; the method selects points that have a small condition number with respect to translations, and consequently, to rotations [8]. It is also worth mentioning that f_1, $\sqrt[n]{f_2}$, f_3, and f_4 are equivalent modulo the choice of a convenient matrix norm [9].

4. RESULTS AND DISCUSSION

The detectors described above have been evaluated on synthetic[1] and real 3D medical images depicting different anatomical scenes. The dataset comprised magnetic resonance (MR) brain images, wrist and feet computerized tomography (CT) scans. Table 1 summarizes the image modalities and observed anatomical scenes used in the experimental setup. The detectors were implemented with integration and derivation scales[2], σ_I and σ_D, respectively, set to 1.5; ϵ, the constant for f_1, was set to 0.05 and, for f_4, p was set to 3. We have explicitly neglected the threshold given in (4) in order to yield more distinctive results for f_1 and f_2. Preliminary experiments have shown that the removal of unsignificant points (pure edge points) via (4) has a tendency to produce less distinctive patterns and less repeatable points as well, since edge points

[1] Available at `http://www.bic.mni.mcgill.ca/brainweb/`.
[2] A Gaussian kernel was used as the weighting function.

are removed and they tend to be less prone to distortions conveyed by noise addition or geometric transformations. Herein, unlike most of the generic evaluations of detectors, the repeatability rate is not directly assessed; our primordial objective is to analyze the behavior of a point based registration technique in which the studied local features are given as landmarks. The chosen registration algorithm was the Coherent Point Drift (CPD) method, a probabilistic solution that can deal with non-linear non-rigid transformations. For a detailed description of the method, we refer to [10]. Fig. 1 illustrates the different detection results on a slice of a human brain MR volume used in the experiments. One can see that results are slightly different. Although some landmarks appear to be identified by more than one detector, the repeatability rate computed among the sets of 500 landmarks, in a 1.5-neighborhood, shows that results are diverse as outlined in Table 2, where r is the repeatability rate function [11] and V_{f_i}, $i = 1, \ldots, 4$, denotes the set of landmarks detected by f_i.

Set	Description	Size	Remarks
D1	6 brain MR images	$181 \times 217 \times 181$	Synthetic images T1-, T2-weighted 1 mm slice thickness 0%,3% and 5% (noise levels)
D2	1 brain MR image	$448 \times 512 \times 20$	T1-weighted 3 mm slice thickness
D3	1 knee MR image	$512 \times 512 \times 24$	PD-weighted 4 mm slice thickness
D4	1 wrist CT scan	$512 \times 512 \times 27$	2 mm slice thickness
D5	1 feet CT scan	$512 \times 512 \times 30$	1 mm slice thickness 30 first slices out of 250

Table 1. Dataset for the experiments outlined in Section 4.

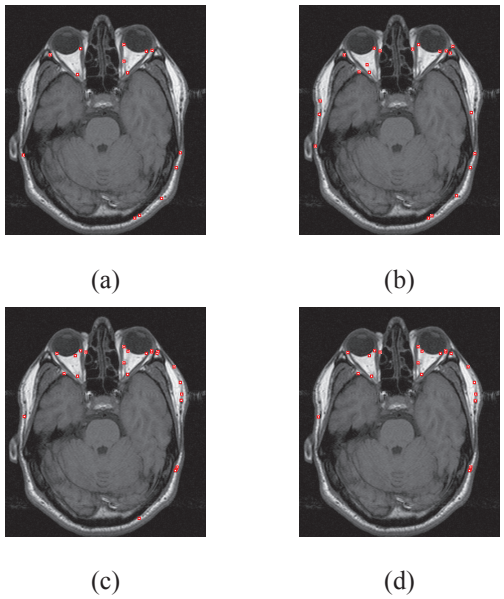

Fig. 1. Results of landmark detection for a MR image of a human brain (slice 12) using the following interest point detectors: (a) f_1; (b) f_2; (c) f_3; (d) f_4.

$r(\cdot,\cdot,1.5)$	V_{f_2}	V_{f_3}	V_{f_4}
V_{f_1}	69.4%	49.6%	56.6%
V_{f_2}		29.6%	34.6%
V_{f_3}			88.6%

Table 2. Repeatability rate results among the sets of interest points detected on the test image D2.

4.1. Distinctiveness Evaluation

Distinctiveness has been assessed by measuring the information content conveyed by the results of detection, a procedure analogous to the one presented in [11] where this criterion was introduced. The goal is to analyze the patterns at interest points and see how distinctive they are. To describe the neighboorhood of interest points, we have computed the 3D-SIFT descriptors [12] at each interest point location, providing an invariant representation of the region. The information content is the entropy exhibited by these invariants for a set of points. In the experiment, 400 interest points, uniformly distributed, were randomly selected over each image. Tables 3 and 4 summarize the results for the distinctiveness of the detectors. For the set D1 and the modalities T1, T2 and CT, we present the average information content.

Dataset	f_1	f_2	f_3	f_4
D1	0.49	0.51	0.5	0.4
D2	0.6643	0.6589	0.8895	0.6339
D3	0.3589	0.3671	0.3256	0.3648
D4	0.6740	0.6874	0.6749	0.6863
D5	0.4528	0.4623	0.4249	0.4646

Table 3. Results of distinctiveness in terms of information content for different modality datasets and detectors.

Modality	f_1	f_2	f_3	f_4
T1	0.59	0.56	0.69	0.57
T2	0.41	0.5	0.37	0.29
PD	0.36	0.37	0.33	0.36
CT	0.56	0.57	0.55	0.58

Table 4. Results of distinctiveness in terms of information content for different anatomical datasets and detectors.

4.2. Non-Rigid Interest Point Based Registration Evaluation

To analyze interest point based registration results and, by inference, conclude on the performance of each of the detectors under a MIR context, the evaluation has to focus on registration results. However, it should reflect the performance of the detectors rather than the registration technique. Therefore, by adopting the CPD algorithm, we have decided to set the parameters required by the method not according to the given set of points, but solely based on the transformation. The analysis has been performed by identifying 500 interest points on each test image. The landmarks sets initially detected were defined as the reference point sets. For each pair of set of points, the reference and the template set, corresponding, respectively, to interest points detected on the original image and on its transformed version, a registration of the template set, T_R, onto the reference set, R, has been performed. Results were analyzed

by computing $r(R, T_R, 1.5)$, the repeatability rate between the given sets in a 1.5-neighborhood. Concerning the transformations, we have applied either small arbitrary rotations to the volumes along the 3 axis (min. of 0^0, max. of 3^0) or added Gaussian noise ($\mu = 0$, $\sigma^2 = \{0.00001, 0.00003\}$) to each axial slice of the test images. For any of the rotated volumes that we have registered onto the original ones, we have set β, the parameter that controls the smoothness of the transfomation, to 2. When dealing with noisy images, β has been lowered to 0.1. The discrepant values were selected in order to yield local smooth transformations in the case of noise and rigid ones in the case of rotation. In all experiments, the third parameter from the registration method, σ, which represents the initial standard deviation in the Gaussian Mixture Model that is fitted to one of the sets of points, was set to 1. Tables 5 and 6 outline the results of registration in terms of repeatability rate applied to rotated and noisy images after 100 and 200 iterations. For space reasons, we will not display all the results, although they will be discussed.

Rot. (deg.)			**Set**	100 iter.				200 iter.			
x	y	z		f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
0	0	3	**D1**	88.9%	85.6%	77.4%	85.7%	89.1%	89%	84.8%	86.9%
			D2	39.4%	46%	35.4%	36.2%	93.2%	92%	87.6%	91.4%
			D3	78%	76%	74.6%	77%	93.4%	92.8%	88.8%	91.8%
			D4	41.6%	38.6%	48.2%	47.6%	83%	81.6%	79%	83%
			D5	45.6%	33.2%	27.8%	28.2%	89%	89.2%	84.4%	86.6%

Table 5. Results in terms of repeatability rate for registered rotated volumes (average rates for the set D1).

Gaussian noise		**Set**	100 iter.				200 iter.			
μ	σ^l		f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
0	0.00001	**D2**	5.6%	5%	6%	9.4%	88.8%	90.4%	88.6%	91.4%
		D3	11.6%	12%	11.4%	14.2%	95.4%	93.2%	92.6%	98.4%
		D4	8.2%	7.6%	9.8%	12.4%	91%	87.4%	93.8%	99%
		D5	5.6%	5.8%	4.6%	7.2%	93.4%	92%	92.2%	94.6%

Table 6. Results in terms of repeatability rate for registered noisy volumes.

We have performed an additional set of experiments on synthetic images that were generated with different levels of noise calculated relative to the brightest tissue. Table 7 summarizes the repeatability rate results after a registration of the noisy images onto the noise free ones. The number of iterations was again set to 200.

Noise	**Modality**	f_1	f_2	f_3	f_4
3%	**T1**	93.6%	93.4%	92.8%	100%
	T2	94.6%	93.4%	91.6%	100%
5%	**T1**	93.4%	94.2%	93.2%	100%
	T2	93.6%	90.8%	93.2%	100%

Table 7. Results in terms of repeatability rate for registered noisy synthetic volumes.

4.3. Discussion

From the observation of the results of the experiments described in 4.1 and 4.2, an important conclusion to be drawn is that finding an appropriate detector to support MIR in such a context as the one depicted herein is still an open issue. The suitability exhibited by f_4 can be enhanced if we rely also on a reliable local description, although it will imply a considerable increase of computational complexity. Furthermore, the detector requires 200 iterations in the CPD algorithm to yield better results. f_1 and f_2 tend to exhibit a similar performance and have shown to be stable in the presence of noise and image rotations, however if we apply the suggested threshold, their performance could decrease. f_3 exhibits the most heterogeneous results in terms of performance; one elucidative example is the different level of distinctiveness exhibited for the T1 and T2-weighted tomographies from our dataset.

5. CONCLUSIONS

We have performed a comparative study on interest point detectors under a medical image registration context, relying on gradient-based solutions known for their robustness to common image distortions and a low computational complexitty as well. Our experimental setup comprised several modalities of medical images depicting different anatomical scenes. The analysis has relied on the registration results supported by the feature detectors and on their distinctiveness. Results have shown a slight outperformance of the Kenney *et al.* detector, which is more suitable to deal with noisy images. In addition, the method exhibits more homogeneous results regarding both criteria when compared to the remaining ones.

6. REFERENCES

[1] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey," *Found. Trends Comput. Graph. Vis.*, vol. 3, no. 3, pp. 177–280, 2008.

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

[3] A. Noble, *Descriptions of image surfaces*, Ph.D. thesis, Department of Engineering Science, University of Oxford, 1989.

[4] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. of the 4th ALVEY Vision Conference*, 1988, pp. 147–151.

[5] K. Rohr, "On 3d differential operators for detecting point landmarks," *Image Vision Comput.*, vol. 15, no. 3, pp. 219–233, 1997.

[6] J. Shi and C. Tomasi, "Good features to track," in *Proc. of the 1994 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94)*, 1994.

[7] C. S. Kenney, B. S. Manjunath, M. Zuliani, G. Hewer, and A. Van Nevel, "A condition number for point matching with application to registration and post-registration error estimation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1437–1454, November 2003.

[8] M. Zuliani, C. Kenney, and B. S. Manjunath, "A mathematical comparison of point detectors," in *Second IEEE Image and Video Registration Workshop (IVR)*, Jun 2004.

[9] C. S. Kenney, M. Zuliani, and B.S. Manjunath, "An Axiomatic Approach to Corner Detection," in *Proc. of the 2005 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 1, pp. 191–197.

[10] A. Myronenko, X. Song, and M. Carreira-Perpinan, "Non-rigid point set registration: Coherent point drift," in *Advances in Neural Information Processing Systems (NIPS) 19*, B. Schölkopf, J. Platt, and T. Hoffman, Eds., pp. 1009–1016. MIT Press, Cambridge, MA, 2007.

[11] C. Schmid, R. Mohr, and C. Bauckhage, "Comparing and evaluating interest points," in *Proc. of the Sixth Int. Conf. on Computer Vision (ICCV)*, 1998, pp. 230–235.

[12] P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional sift descriptor and its application to action recognition," in *MULTIMEDIA '07: Proc. of the 15th International Conference on Multimedia*, New York, NY, USA, 2007, pp. 357–360, ACM.