Comparison of Salient Point Detection Methods for 3D Medical Images

Bryn A. Lloyd[1,2,3], Gábor Székely[1], Ron Kikinis[3], and Simon K. Warfield[2]

[1] Computer Vision Laboratory, Swiss Federal Institute of Technology, Zürich, blloyd@ee.ethz.ch,
[2] Computational Radiology Laboratory, Children's Hospital, Harvard Medical School, Boston.
[3] Surgical Planning Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston.

Abstract. Salient points are used for various applications, such as medical image registration, tracking, stereoscopic matching. The purpose of this paper is to compare two commonly used methods to extract salient points in 3D medical images. We give an interpretation of the methods and validate their performance empirically based on criteria derived for the task of image registration, displacement measurement and tracking in medical images.

1 Introduction

There has been considerable research on finding an optimal salient point detector. These detectors are often called corner detectors, because they have been designed to find grey-value corners. Typically there aren't many real corners in medical images, such as MR images of the brain. Nevertheless, it is useful in many registration and tracking methods to extract points with high intensity variations in all directions, because they may be matched more reliably. An illustrative example is a straight line, which has high variation perpendicular to the axis, but no variation along the axis, effectively preventing localization in this direction. Therefore the salient point detector should detect points that are distinctive in their environment. Furthermore the detector should be invariant under certain (geometric and photometric) transformations. Mostly research has been done on 2D salient point detectors. Many approaches have been proposed in the literature. Some of the main directions have been:

1. methods based on first extracting the edges as chain codes and then finding points on the edges with maximal curvature[1].
2. differential methods based on directly finding points from the grey-level image. These methods are usually based on calculating derivatives of the image and using some kind of cornerness measure.
3. scale-space methods based on computing features at multiple scales (*e.g.* after wavelet transforming the image [2]).

The first type depends on the accuracy and robustness of the edge detection process, which is notoriously unreliable for images with low signal to noise ratio. For our purposes we prefer the second type of corner detectors. In 2D the most prominent operators of this type, are the the DET operator, which is the determinant of the Hessian matrix, proposed by Beaudet [3], the Kitchen and Rosenfeld [4] operator, which is the curvature multiplied by the gradient magnitude and the Plessey corner detector by Harris and Stephans [5]. One of the short comings of this type of salient point detectors, are the inaccurate localization of corners. Many more recent approaches, such as the Curvature Scale Space method [6], are improvements of these basic operators.

More recently methods to extract points in 3D volume images have been developed. A generalization of some of the operators mentioned above to 3D was given by Rohr [7]. Ruiz-Alzola et al. [8] extended a modified version of the Harris corner detector, by introducing generalized correlation matrices, to deal with vector and tensor pixel values in arbitrary image dimensions.

Naturally the quality of a salient point detector is assessed under certain criteria, that may vary between applications. Typical criteria are a) the invariance of the method under geometric and photometric changes, or over different scales (scale space), b) the accuracy of localization of the points and c) the distinctiveness of the detected points in their local environment. In our setting we would like to use the detected points for registration of medical images from different image modalities. We have defined some properties, which a salient point detector should have for this type of application. For each of these properties we have found a metric to quantify the performance of the operator. In Section 2 we describe two interesting salient point detectors. In Section 3 we discuss several metrics to validate the performance of the detectors. Finally 4 gives some results on the performance of the validated detectors. All of the discussed detectors have been implemented as open source by using and extending the Insight Segmentation and Registration Toolkit [9].

2 Salient Point Detection Methods

In this paper we compare two different methods to extract salient points. The first type of method, is based on finding extrema of the Gaussian curvature. Beaudet proposed a 2D operator and called it DET, as it is the determinant of the Hessian matrix. The usual interpretation is based on the observation that DET is the numerator of the Gaussian curvature:

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{1 + f_x^2 + f_y^2} = \frac{DET}{1 + f_x^2 + f_y^2} \tag{1}$$

The Gaussian curvature is the product of the two principal curvatures. The principal curvatures are the minimum and maximum values of the local curvature of a surface(*i.e.* the image intensities). A positive Gaussian curvature value means the surface is locally either a peak or a valley. A negative value means the surface

locally is a saddle point. And a zero value means the surface is flat in at least one direction. The local extrema of the curvature are selected as points of interest.

The extension to 3D is mainly due to the work of Monga et al. [10], who showed it is possible to calculate locally the Gaussian curvature without fitting a surface. The estimated curvature is the curvature of an isointensity surface in the volume image.

$$K = \frac{f_x^2 \left(f_{yy} f_{zz} - f_{yz}^2\right) + 2 f_x f_{xz} \left(f_y f_{yz} - f_z f_{yy}\right) + cycl.\,(x,y,z)}{f_x^2 + f_y^2 + f_z^2} \tag{2}$$

Where $cycl.(x,y,z)$ stands for cyclic permutation of the coordinates. Thirion and Gourdon [11] used this expression to find crest lines in 3D CT and MRI images. The same expression was obtained by Florack [12] by applying invariance theory.

The second type of operator in our comparison is based on the corner detector proposed by Harris and Stephans [5]. An estimate of the autocorrelation function is calculated at each pixel location using the first derivatives of the image intensities. An Eigenanalysis of the estimated correlation matrix provides information about how the intensities are changing locally. The method was extended to 3D by Rohr [13]. He also gives a sound statistical interpretation to the correlation matrix and its principal invariants. Finally Ruiz-Alzola et al. [8] generalized this method to vector and tensor valued images by introducing generalized correlation matrices, which are the average over the correlation matrices of each vector/tensor component. For scalar pixels the correlation matrix C_g is given by the expectation over the dyadic product of the 3D intensity gradient.

$$\mathbf{C}_g = E\left[\mathbf{g(x)}\ \mathbf{g(x)}^T\right] \tag{3}$$

Usually the expectation is approximated by the mean in a neighborhood around the pixel. If all Eigenvalues of C_g are large, a point has a lot of structure. If one or more Eigenvalues are small, there is not much change of intensities in one or more directions. For salient points the determinant, which is the product of the Eigenvalues, should therefore be large. The trace of the correlation matrix corresponds to the mean square gradient norm. It should be large along edges and at salient points. Most cornerness measures are ratios or linear combinations of the determinant and trace. ($e.g.\ det\,(C_g)\,, \frac{det(C_g)}{tr(C_g)}, det\,(C_g) + ktr\,(C_g)$).

A statistical interpretation of the estimated correlation matrix can be provided using the Cramer-Rao bound, which relates the covariance matrix to the correlation matrix:

$$\Sigma_g = \frac{\sigma_n^2}{m} \mathbf{C}_g^{-1} \tag{4}$$

The Eigenvalues of Σ_g correspond to the localization uncertainty in the direction of the associated Eigenvector. The ellipsoid described by the Eigenvectors and Eigenvalues can be interpreted as the error ellipsoid. The salient point detector should select points with a small volume of the ellipsoid. Rohr uses the principal invariants of Σ_g as a cornerness measure and justifies each of these measures based on minimizing the error ellipsoid.

In this study we chose $det\,(C_g)$ which as a cornerness measure, which was proposed by Rohr [13]. To extract the salient points, we calculate the local maxima of the cornerness measure. In both methods thresholding can be applied to reduce the number of points with low responses to the operator.

3 Validation Methods

Our approach is task driven. The points are typically used to find corresponding points in two images. The correspondences can be used for landmark based non-rigid image registration techniques and motion tracking. Various authors have validated point detectors, usually for synthetic date, *i.e.* real corners. We believe there is a need to improve methods to empirically validate existing salient point detectors. These validation methods should reflect the purpose of the point detector and help choosing a method for a given application. Ultimately good validation methods could also be used to guide parameter selection for the chosen detector. Of course validation metrics are hard to choose and reflect only certain aspects. For if a validation metric would reflect the gold standard, we should simply use it as the point detector. The solution might be to use several metrics. Therefore we define the following criteria to compare the quality of the detection methods:

1. **Distinctiveness** and information content. Sebe et al. [2]) calculate the entropy of the intensities in a neighborhood of the salient point.
2. **Invariance** under geometric and photometric(linear) transformations. These properties have been analyzed theoretically by several authors. Both detectors in this study are theoretically invariant to rigid transformations(*e.g.* [13], [11]). It is easy to see that both detectors are relative invariant to linear intensity changes (i.e. $i'_k = s_k i_k + u_k$). Therefore we have not defined a metric for this criterion.
3. **Robust to scale changes** (*e.g.* Sebe et al. [2]) define a repeatability rate, which counts how many points are found at two different scales. We calculate the repeatability rate by applying different amounts of smoothing (variance 1.0, 2.0 and 3.0). A point is considered to be present in both images, if there is a corresponding point in its predefined neighborhood. We selected the neighborhood radius as 3 mm, which corresponds to about 1–2 voxels in our data.
4. **Repeatability** between registered images. The detector should find "true" salient points. True salient points should be possible to find, even in images scanned at different times. Again we define a repeatability rate, which is calculated between registered images. The images are rigidly registered using a multi-resolution method and a mutual information similarity metric.

4 Results

We performed this study using thirteen MR images of the brain. Eight MPRAGE images and one T2 weighted image were obtained in a 1.5T scanner. Four SPGR

images were also obtained. The MPRAGE images have a resolution of 0.93mm x 0.93mm x 1.3mm. The SPGR images are of lower quality, acquired in a 0.5T scanner and have a resolution of 0.86mm x 0.86mm x 2.5mm. We implemented both methods as open source using and extending the Insight Toolkit [9][4].

The image intensities are first rescaled between 0 and 10. All derivatives are computed by convolving the image with a derivative of a Gaussian, were the Gaussian variance was set to 1.0. The correlation matrix Cg was estimated using Expression 3, were the expectation was replaced by the mean operator in a 3x3x3 window. The Gaussian curvature was computed using first and second order derivatives according to Expression 2. Local maxima were extracted by selecting points that are larger than all neighbors in a window of 3x3x3 voxels. For anisotropic pixel spacings we modified the window sizes accordingly. For the correlation method we selected a threshold of 2.0. For the curvature method we chose a threshold of 0.5. These thresholds depend of course on the intensity rescaling and were selected in a way that a comparable number of points results for both methods. The selected thresholds were set low, which is why we detected more than 10000 points for a full volume MR image of the brain.

Figure 1 shows a slice of a 3D volume image. The image on the left shows the typical points, which are picked up by the curvature method. The image on the right shows the method based on the estimate of the correlation matrix. The methods typically find different types of salient points. One difference that we observed is that the curvature method finds more points by the ventricles. Table 1 shows some summary statistics of the chosen metrics. In the first column the average entropy of the selected salient points is shown. Using the Pearson correlation coefficient we found the curvature method to be more correlated to the entropy in a window of 7x7x7 voxels. The entropy cannot be a complete validation metric though, since it incorporates no spacial information. Both methods achieve about the same average entropy.

Method	entropy	repeatability rate at different scales (13 cases)	repeatability rate for registered images (same modality, 1 case)	repeatability rate for registered images of different modality (3 cases)
Curvature	0.528	0.68 ± 0.10	0.69	0.30
Correlation	0.532	0.78 ± 0.03	0.76	0.28

Table 1. Summary statistics for our comparison. The shown values are the mean values obtained. The entropy is similar for both methods. The repeatability rate for images at different scales, is higher for the correlation method. For the registered images of the same modality (MPRAGE), the repeatability rate is also higher for the correlation method, whereas it is similar and low for registered images of different modality (MPRAGE and SPGR).

[4] The code is being developed in the NAMIC SandBox SVN repository, and can be checked out using the following command: `svn checkout http://www.na-mic.org:8000/svn/NAMICSandBox/PointLandmarkDetection`

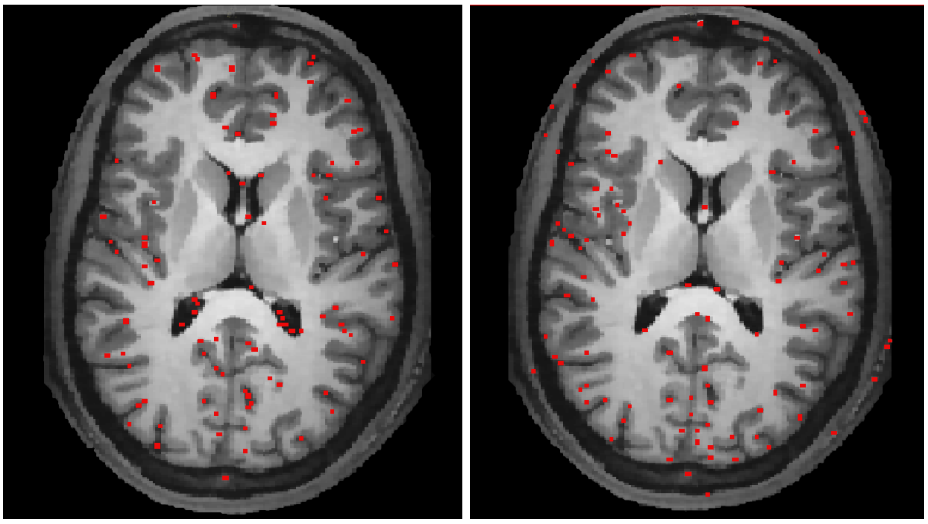

Fig. 1. This figure shows detected salient points in a axial slice of a brain. In a) Beaudet/Thirion curvature based detector, in b) the Harris/Rohr correlation based method is shown.

The repeatability rates were calculated using the 6000 points with the highest cornerness value in the image. In the case of the scale changes, the points were selected from one image by applying the method with different amounts of smoothing (variance 1.0, 2.0 and 3.0). The repeatability rates were higher for the Rohr correlation operator. The obtained rates are comparable to the repeatability rates obtained by Sebe et al. [2]), although their experimental setup was quite different. The repeatability rate for registered images was calculated for four image pairs. Three of these pairs were images of different modality (MPRAGE and SPGR). For these image pairs the high resolution images were resampled to the same voxel spacing as the images of lower resolution. The repeatability rate for registered images of the same modality were also higher for the Rohr operator. For images of different modality the repeatability rates were poor for both methods.

5 Discussion

In this short paper we have described two different methods to extract salient points in 3D images. Both methods use differentials of the image intensities to calculate a measure of saliency. In our study we found it difficult to validate methods to extract landmarks by visual inspection, due to the number of extracted points per image. We validated the methods using several metrics, which were derived from task specific criteria. The validation was carried out using MR images of the brain. The method using the Rohr operator, which is the determinant of the correlation matrix, had higher repeatability rates w.r.t scale changes

(smoothing). Both methods had a low repeatability rate for registered image of different modality.

6 Acknowledgements

This investigation was supported in part by NSF ITR 0426558, a research grant from the Whitaker Foundation, a research grant from CIMIT, grant RG 3478A2/2 from the NMSS, and by NIH grants R21 MH67054, R01 LM007861 and P41 RR13218.

References

1. Medioni, G., Yasumuto, Y.: Corner detection and curve representation using cubic b-spline. IEEE int Conf., Robotics and Automation (1986) 764–769
2. Sebe, N., Lew, M.S.: Comparing salient point detectors. Pattern Recognition Letters **24** (2003) 89–96
3. Beaudet, P.R.: Comparing slient point detectors. In: International Joint Conference on Pattern Recognition. (1978) 579–583
4. Kitchen, L., Rosenfeld, A.: Gray-level corner detection. Pattern Recognition Letters **1** (1982) 95–102
5. Harris, C., Stephans, M.: A combined corner and edge detector. (1988) 147–151 4th Alvey Machine Vision Conference.
6. Mokhtarian, F., Suomela, R.: Robust image corner detection through curvature scale space. IEEE Trans. on Pattern Analysis and Machine Intelligence **20** (1998) 1376–1381
7. Rohr, K.: On 3d differential oparators for detecting point landmarks. Image and Vision Computing **15** (1997) 219–233
8. Ruiz-Alzola, J., Kikinis, R., Westin, C.F.: Detection of point landmarks in multidimensional tensor data. Signal Processing **81** (2001) 2243–2247
9. : (Insight segmentation and registration toolkit) http://www.itk.org/.
10. Monga, O., Benayoun, S.: Using partial derivatives of 3d images to extract typical surface features. Computer Vision and Image Understanding **61** (1995) 171–189
11. Thirion, J.P., Gourdon, A.: Computing the differential characteristics of isointensity surfaces. Computer Vision and Image Understanding **61** (1995) 190–202
12. Florack, L.: The syntactical structure of scalar images. PhD thesis (1993)
13. Rohr, K.: Extraction of 3d anatomical landmarks based on invariance principles. Pattern Recognition **32** (1999) 3–15