Computerized Medical Imaging and Graphics 36 (2012) 634–642

Contents lists available at SciVerse ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Cervical spine mobility analysis on radiographs: A fully automatic approach

Fabian Lecron*, Mohammed Benjelloun, Saïd Mahmoudi

University of Mons, Place du Parc, 20, 7000 Mons, Belgium

ARTICLE INFO

Article history:
Received 5 May 2012
Received in revised form 19 July 2012
Accepted 23 August 2012

Keywords:
Cervical spine motion
Vertebra detection
Vertebra segmentation
X-ray images

ABSTRACT

Conventional X-ray radiography remains nowadays the most common method to analyze spinal mobility in two dimensions. Therefore, the objective of this paper is to develop a framework dedicated to the fully automatic cervical spine mobility analysis on X-ray images. To this aim, we propose an approach based on three main steps: fully automatic vertebra detection, vertebra segmentation and angular measurement. The accuracy of the method was assessed for a total of 245 vertebræ. For the vertebra detection, we proposed an adapted version of two descriptors, namely Scale-invariant Feature Transform (SIFT) and Speeded-up Robust Features (SURF), coupled with a multi-class Support Vector Machine (SVM) classifier. Vertebræ are successfully detected in 89.8% of cases and it is demonstrated that SURF slightly outperforms SIFT. The Active Shape Model approach was considered as a segmentation procedure. We observed that a statistical shape model specific to the vertebral level improves the results. Angular errors of cervical spine mobility are presented. We showed that these errors remain within the inter-operator variability of the reference method.

1. Introduction

A wide majority of adults are, or will be, confronted to back problems in their daily life. For instance, low back pain is a leading cause of work absenteeism. Another illustration concerns the cervical spine trauma which represents the majority of spinal lesions. In this context, whiplash injuries are probably the most recurrent cause for insurance claims but with no well-established diagnosis [1].

Many of these problems are the consequence of an abnormal spinal motion. One practical device to extract spine motion in real-time is the digital videofluoroscopy (DVF). Proposed for the first time in [2], this technique consists in recording X-ray images as video frame images. At that time, the authors showed that it was possible to acquire and study the lumbar spine motion with such a system.

The initial step for studying spine motion in a DVF system is landmarking or locating the vertebræ among the frames. Initially, this stage was operated manually. Nowadays, the contributions in the literature attempt to automate as much as possible the landmarking or vertebra locating procedures. Common approaches are based on template matching. The idea is to build a template defining a vertebra, to place the template on the first frame of the video sequence (generally, this step requires a human intervention) and finally to track the template frame by frame. In this context, approaches based on cross-correlation are presented in the literature. The idea is to define a measure of correlation which can be regarded as a *similarity* value. In [3], Bifulco et al. used the cross-correlation for automatic recognition of vertebra landmarks. Muggleton and Allen presented a similar approach in [4], but refined it with a second stage where an annular template enclosing the vertebral body edges is used. The cross-correlation has also been considered by Cerciello et al. in [1]. Recently, they proposed an improved version of template matching in [5]. The vertebra edges are enhanced by estimating image gradient. However, they made the observation that fluoroscopic images are very noisy and that gradient estimation is precisely sensitive to noise. Therefore, this issue is solved by a noise suppression technique preserving edges.

Other methods aimed at extracting vertebral boundaries and tracking them in a video sequence are presented. In this context, Zheng et al. integrated Fourier descriptors with Hough transform algorithm in [6]. They obtained promising results on a calibration model but the approach was validated on only 2 real sequences. Another interesting method was developed in [7], where Lam et al. presented a tracking algorithm based on a dynamic Bayesian network with a particle filter at each node. In this framework, prior information of kinematics and anatomical configuration are combined with the information collected from each frame to estimate the present state. Finally, Reinartz et al. argued in [8] that exact contour extraction is not necessary and proposed a method based on the normalized gradient field for tracking vertebræ.

All these contributions in the context of DVF are of great interest. However, conventional X-ray radiography is the most common modality used in emergency rooms since it is relatively

* Corresponding author. Tel.: +32 65 37 40 58.
E-mail address: Fabian.Lecron@umons.ac.be (F. Lecron).

0895-6111/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compmedimag.2012.08.004

inexpensive and fast. Moreover, it remains nowadays the most common method to analyze spinal motion in two dimensions [7]. Surprisingly, the literature on the subject does not focus on automating approaches. Computer-aided protocols such as [9] or [10] require the intervention of an operator on the images. A semi-automatic approach is proposed in [11].

Actually, works on fully automatic vertebra extraction on radiographs are few in the literature. Zamora et al. tried to take advantage of the Generalized Hough Transform (GHT) in [12] but they present a segmentation rate equal to 47% for lumbar vertebræ without providing information about the detection rate. Recently, Dong and Zheng [13] have proposed a method combining GHT and the minimal intervention of a user with only 2 clicks in the image. A fully automatic approach has been developed by Casciaro and Massoptier in [14]. They use a shape constraint characterization by looking for every shape that could be an inter-vertebral disc. Very recently, we presented in [15] an approach to locate the anterior vertebra corners in a radiograph. Nevertheless, these landmarks are not sufficient in order to analyze the cervical spine motion. Therefore, this approach can be combined to a segmentation procedure.

Our contribution is to propose a fully automatic framework for cervical spine mobility analysis on radiographs. Our framework is based on three main steps: fully automatic vertebra detection, vertebra segmentation and vertebra mobility analysis. More precisely, the vertebræ on a given radiograph are detected by the anterior corners. This information is used to initialize a segmentation procedure based on Active Shape Model [16]. Once the exact contour of the vertebræ is extracted, a protocol for cervical spine mobility analysis on radiographs can be performed. To the best of our knowledge, this work is the first report on the subject, since only a few semi-automatic methods have been proposed.

The rest of the paper is organized as follows. Section 2 presents the approach for fully automatic vertebra detection on radiographs. Some aspects of our previous work [15] are adapted to deal with mobility analysis. In Section 3, we describe how the vertebra detection is used to initialize a segmentation procedure based on Active Shape Model. Section 4 reports the detection, segmentation and mobility results for our framework. Section 5 concludes the paper.

2. Vertebra detection

The vertebra detection is operated with a sequence of 3 stages: interest point detection, interest point description and interest point classification.

2.1. Interest point detection

Medical images such as radiographs present large homogeneous areas making classical detection algorithms inefficient. To detect keypoints, we propose to use a method based on the polygonal approximation of the edges [17]. Therefore, we consider as keypoint any point at the intersection of two segment lines. First, the contours in the image are detected with a classical filter. Then, a polygonal approximation of the resulting edges is performed. Points located at the intersection of two segment lines are retrieved. Depending on the application and on the object of interest, an angle filter could be carried out at this step. If, in a given application, a keypoint needs to be defined as any point at the intersection of two segment lines defining a right angle, an angle filter could be used to decide which points are at the intersection of two segment lines with a right angle. Nevertheless, we have not used an angle filter in this work.

Our approach for detecting keypoint is obviously scale, translation and rotation invariant. The detection invariance means that if the image is rotated, if a greater or a lower patient is considered, our approach of detection based on polygonal edge approximation will still be able to detect keypoints in the image.

2.2. Interest point description

All the detected points have now to be *described*. For this stage, we decided to reuse some aspects of Scale-invariant Feature Transform (SIFT) [18] and Speeded-up Robust Features (SURF) [19] algorithms.

2.2.1. SIFT

The advantage of this descriptor is the scale, translation and rotation invariance and the partial invariance to illumination changes. The description invariance means that the features of a keypoint remain valid if the image is scaled, rotated or translated.

The first stage of the algorithm consists in detecting local extreme of difference-of-Gaussian filter at different scales in the image. Therefore, in our case, the keypoints are already detected in the image. As a consequence, the next step is to compute an orientation histogram from the orientations of the points close to the pixel of interest. The highest peak detected in the histogram stands for the orientation of the interest point.

Afterwards, we can rotate the coordinates of the point with an angle equal to the orientation. Then, a 4×4 window is created around the pixel. The 16 resulting regions are also divided into smaller ones. Inside them, the gradient magnitude of every point is computed and an orientation histogram discretized in 8 classes is built. The quantity added in the histogram corresponds to the gradient magnitude at the current point. Nevertheless, pixels far from the point of interest must have a smaller influence than close ones. For this reason, a Gaussian function, with a σ equal to 1.5 times the descriptor window size, is used to weight the histogram.

The orientation histogram is discretized in 8 classes. Since the window is divided into 16 regions, the descriptor is characterized by 128 features. Experimental tests in original Lowe's paper [18] showed that there is significant difference in accuracy if the orientation histogram is discretized in more than or less than 8 classes. We refer the reader to [18] for more detail on these experimental tests.

2.2.2. SURF

First of all, a 4×4 grid has to be placed around the interest point. The size of the grid is fixed by the user. This variable constitutes a parameter and will be studied in Section 4. The scale s of the point can be deducted with the size of the descriptor window. The size of the window is equal to $20s$. The following steps are equivalent to the description steps depicted in [19]. The orientation of the point is computed. Then, the 4×4 window is rotated according to this orientation. Finally, the keypoint is defined with a vector of 64 coordinates.

2.3. Interest point classification

In order to detect a specific object, we need to classify the detected points and answer the question: *"Do this keypoint belong to an object of interest?"*.

To this aim, we propose to use the Support Vector Machine (SVM) approach. Introduced by Vapnik [20], SVM is widely dedicated to classification and regression problems. Consider that we have a set of data $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^n \times \pm 1$. The goal of SVM is to determine the hyperplane separating at the best the data into 2 distinctive classes: +1 and −1. The hyperplane $\mathcal{H}$ is defined by the couple (w, b) such as $\langle w, \Phi(x) \rangle + b = 0$, where $\Phi(x)$ transposes the data in a higher-dimensional space $\mathcal{F}$, called *feature space*.

Actually, w and b are obtained by solving the quadratic programming problem (1).

$$\begin{cases} \underset{w,b}{\text{minimize}} & \frac{1}{2}\|w\|^2 \\ \text{subject to :} & y_i(\langle w, \Phi(x_i)\rangle + b) \geq 1, \quad i = 1, \ldots, m. \end{cases} \tag{1}$$

Let us introduce the dual formulation of the problem (1). It is formalized by:

$$\begin{cases} \underset{\alpha}{\text{maximize}} & \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j \langle \Phi(x_i), \Phi(x_j)\rangle \\ \text{subject to :} & \alpha_i \geq 0, \quad i = 1, \ldots, m, \\ & \sum_{i=1}^{m}\alpha_i y_i = 0. \end{cases} \tag{2}$$

If we focus on the dual problem (2), one can notice that actually, the function $\Phi(x)$ is not needed explicitly. Only the inner products $\langle \Phi(x_i), \Phi(x_j)\rangle$ are used in this dual formulation. Let K be a symmetric matrix consisting of the inner products $\langle \Phi(x_i), \Phi(x_j)\rangle$. Any positive semidefinite matrix could be used as a kernel matrix K, since the positive semidefiniteness ensures the existence of the function $\Phi(x)$. Let $(\alpha_1^*, \alpha_2^*, \ldots, \alpha_m^*)$ be an optimal solution of the problem (2). It can be shown that $\langle w^*, \Phi(x)\rangle = \sum_{i=1}^{m}\alpha_i^* y_i k(x, x_i)$.

From this point, we know all the parameters to compute the hyperplane $\mathcal{H}$ separating our data. Therefore, any new point $x \in \mathbb{R}^n$ can be classified by determining which side of the hyperplane x belongs to.

Actually, 3 classes are considered with the SVM classifier: anterosuperior corners (class 1), anteroinferior corners (class 2) and other points (class 3).

2.4. Missed vertebra recovery

It is obvious that the previous steps do not allow to retrieve the location of all the vertebræ. Actually, some corners are not detected during the interest point detection process. These non-detected points can therefore not be classified and be recognized as vertebra corners.

To deal with this issue, we propose to take into account an *a priori* on the spine. The spine of a human being is indeed characterized by the sequence:

Superior corner – Inferior corner – Superior corner – . . . – Inferior corner.

The detected points are assessed in order to see if they verify this vertebral sequence. During this process, geometrical information on the vertebræ can be deduced. Every couple of points verifying the vertebral sequence allows to estimate the intra-vertebral or the inter-vertebral distance of the spine.

With this information, a point that has not been detected can be easily recovered. Let us consider Fig. 1 in order to illustrate our purpose. On the left, one can see that the red corner has not been detected. This anteroinferior corner can be recovered by considering the mean intra-vertebral distance from the corresponding anterosuperior corner with the same orientation than the previous vertebra.

3. Vertebra segmentation

The segmentation methods in the literature are dependent on the imaging modality (*e.g.* Computed Tomography, X-ray Images, Magnetic Resonance Images, *etc.*). With regard to X-ray images, the efficient frameworks to segment vertebræ were based on a model. The most common model-based approaches are Active Shape Model (ASM) and Active Appearance Model (AAM), developed by Cootes et al. in, respectively [16,21]. The former is mainly interested in the shape modelization of a given object while the latter integrates the appearance of the object in the model. ASM has been successfully applied to vertebra segmentation in [22–25] and AAM in [23,26,27]. Nevertheless, ASM and AAM contributions in the literature are often semi-automatic, *i.e.* the user has to click several times on the image to approximate the position of the vertebræ, before the segmentation process.

In the previous section, we presented how to detect the location of the vertebræ by the anterosuperior and anteroinferior corners. With this information, we can propose a fully automatic segmentation procedure. Since we are only interested in the shape information, we decided to work with the Active Shape Model.

3.1. Active Shape Model

Active Shape Model approach is divided into two components: a statistical shape model and a gray level model. The former is used to constrain the shape of the object while the latter guides the segmentation process.

3.1.1. Statistical shape model

Let N be the number of landmarks composing a shape. If we consider a two-dimensional space, a shape is represented by a column-vector x with $J = 2N$ coordinates. A deformable model can be built given a sample of k shapes by computing the mean $\bar{x}$ and the covariance matrix associated to the sample. A principal component analysis (PCA) is then applied to capture the variation in the sample. The PCA yields to a matrix ϕ where each column is an axis (a component) along which the variance is maximized. Therefore, any shape x can be approximated by:

$$x = \bar{x} + \phi d, \tag{3}$$

where d is a vector of PCA weights controlling the deformation of the shape. Actually, these parameters are limited by their variance σ^2.

Eq. (3) represents the statistical shape model. By varying the PCA weights d within the limits σ^2, new shapes can be generated.

3.1.2. Segmentation

Once the statistical model is defined, the mean shape of the model is used to initialize the segmentation. In this work, one mean shape per vertebra has to be positioned. To this end, the framework of vertebra detection presented in Section 2 is applied. Let us remind that this framework is fully automatic.

Now that the vertebra locations are detected, the segmentation process can be performed. To this aim, the gray level variation has to be locally evaluated around each landmark in the sample. Then, a mean profile of the texture (gradient intensity) can be deduced. After the initialization, a local analysis of the texture is carried out around each landmark of the initial shape. The goal is to find the best match with the mean profile previously determined. The distance used for the profile comparison is the Mahalanobis distance. This search implies that the landmarks are moved during the segmentation. The procedure is repeated until the convergence, *i.e.* when the match between the current shape profile and the mean one is no more improved. At each iteration, the current shape has to fit the statistical shape model previously defined.

3.2. Vertebra models

There are several ways to represent vertebræ by a statistical shape model. In this study, only landmarks are used. As a consequence, intra-vertebral distance is implicitly represented by the statistical shape model since it corresponds to the distance between

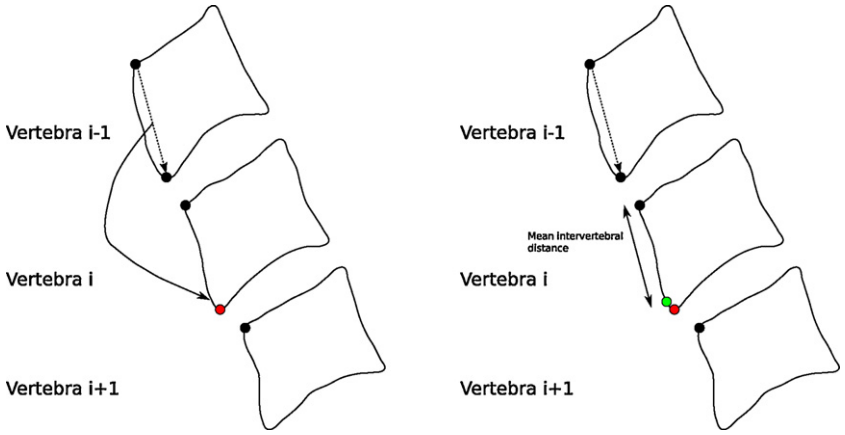

Fig. 1. Recovery of a missed corner.

the first and the last landmarks. We propose to study the influence of two modelizations:

- Model 1: each vertebra is represented by the same model, whatever the vertebral level.
- Model 2: each vertebral level is characterized by a specific model.

If we consider the cervical vertebræ C3–C7, this mean that for Model 1, five identical modelizations are used for the segmentation process whereas five different modelizations are computed in the case of Model 2. Let us remark that, both for Model 1 and Model 2, the five modelizations are independent from each other.

Finally, it is important to clarify that our ASM system does not include a multi-object model. As a consequence, the vertebræ are segmented independently.

4. Experimental results

4.1. Data and methods

The radiographs used in this work were acquired at Jolimont Hospital in Haine-Saint-Paul, Belgium. All the images have been annotated by an experienced radiologist. They were chosen so that all the cervical vertebræ C3–C7 are visible (sometimes partially) with the naked eye. The detection and the segmentation were performed on 245 cervical vertebræ from 49 distinct patients. Thus, the tests are performed on different images and the results are not biased. Patient's age ranges from 35 to 60 years. There are about as many males as females. The radiographs were essentially acquired because of neck pain or neck brachialgia. The radiologist used six points to define one vertebra. The radiographs all present similar resolutions in pixels, since they are focused on the same area of the human body.

In order to assess our framework, we have performed a leave-one-out cross validation based on an execution of the algorithm on all the radiographs. This way, at every run of the algorithm, one image is tested while the rest of the sample is used for the learning phase. As a result, we can ensure a complete independence between learning and testing data. This aspect is also valid for the detection and the segmentation process.

In the following, several measures will be used to assess our framework. One way to measure the segmentation error is to compute the distance between the ASM contour and a theoretical contour defined by an experienced radiologist. Therefore, a standard has been defined for the radiographs of the image sample. The chosen distance for measuring the segmentation error is the point-to-line distance. Used in [25], the principle is to compute the length

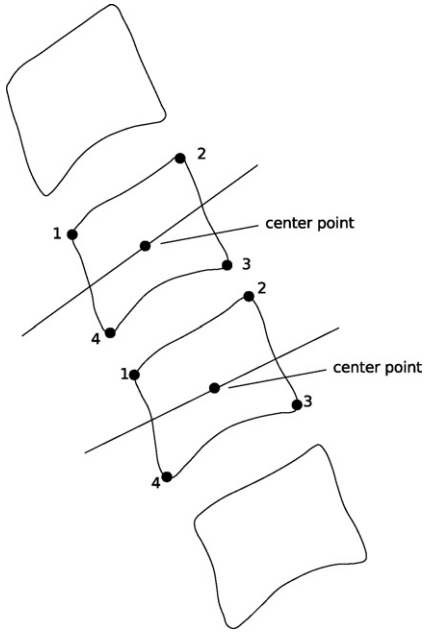

Fig. 2. Protocol for measuring the orientation and the mobility of the vertebræ.

of the perpendicular dropped from each landmark of the theoretical contour to the spline evaluated between the landmarks of the ASM contour.

The vertebra orientation and the vertebra mobility will also be assessed in the further sections. To this end, we used the measurement protocol described in [9]. This protocol is well adapted to our framework since it is based on the computer-aided location of objective landmarks on vertebræ C3–C7.

An illustration of the measurement protocol is shown in Fig. 2. The four vertebra corners are determined by our detection and segmentation framework. The center point is then defined as the geometric center of the four corners. The orientation of a vertebra is determined by the orientation of the vertebra midplane, which is defined as the line running through the midpoints between corners 1, 4 and corners 2, 3, respectively. The motion of a vertebral segment is computed by determining the angle between two adjacent vertebræ. This angle is given by the angle between their midplanes. The angle is counted positive if the wedge opens anteriorly. To sum up, four vertebral segments are considered in this study: C3/C4, C4/C5, C5/C6 and C6/C7. Angles concerning the first vertebra (C3) and the last vertebra (C7) are respectively related to the segments C3/C4 and C6/C7.

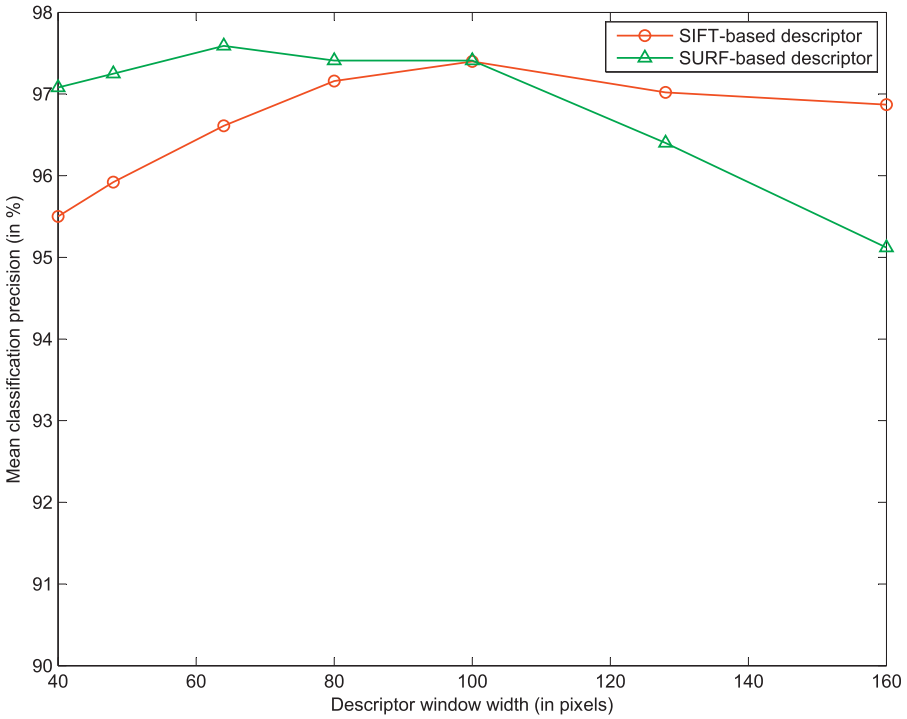

Fig. 3. Evolution of the mean classification precision according to the descriptor window width.

4.2. Performance of vertebra detection

In this section, we assess the performance of two descriptors for the vertebra detection on radiographs. In Section 2.2, we have indeed proposed a modified version of two descriptors: SIFT and SURF. Prevalidation tests allowed us to reveal that the main parameter influencing the results is the size of the descriptor window. Some relevant values can be deduced, based on vertebra characteristics: the intra- and inter-vertebral distances. In our case, the inter-vertebral distance varies between 30 and 40 pixels and the intra-vertebral between 80 and 90 pixels.

We presented in Fig. 3 the evolution of the mean classification precision according to the descriptor window width. The classification precision is calculated and averaged among all the images. A 100% rate indicates that all the points in a given radiograph have been successfully sorted in their associated class. Fig. 3 shows that the descriptor window width evolves from 40 to 160 pixels. For the two descriptors, as long as the descriptor window size is increased, the mean classification precision is improved. However, this improvement is limited to approximately 100 pixels. This means that the radius of the window (equal to 50 pixels) needs to be comprised between the inter-vertebral distance and the intra-vertebral distance.

However, two major differences between the SIFT-based and the SURF-based descriptor appear. First, when the descriptor window width is lower than the limit of 100 pixels, the precision of the SURF-based descriptor is better than the SIFT-based precision. Conversely, when the descriptor window width is greater than the limit of 100 pixels, the SIFT-based descriptor outperforms the SURF-based descriptor whose accuracy is drastically reduced. Despite this, the SURF-based descriptor is characterized by the better mean classification precision.

The corner and vertebra detection rates are presented in Tables 1 and 2. The former shows the results obtained without applying the recovery of missed vertebræ, as described in Section 2.4. The latter is characterized by improved results thanks to the recovery of missed vertebræ. Both for Tables 1 and 2, we used

Table 1
Detection rates without recovery of missed vertebræ.

Vertebra	Corner type	Detection rate			
		SURF-based		SIFT-based	
C3	Sup.	71.7%	62.2%	54.3%	47.8%
	Inf.	80.4%		69.6%	
C4	Sup.	80.4%	66.7%	73.9%	56.5%
	Inf.	84.8%		73.3%	
C5	Sup.	80.0%	68.1%	88.9%	73.3%
	Inf.	82.2%		82.2%	
C6	Sup.	80.0%	59.1%	84.4%	68.9%
	Inf.	71.1%		80.0%	
C7	Sup.	73.0%	16.7%	78.4%	16.2%
	Inf.	16.2%		16.2%	
Mean		72.0%	54.6%	70.1%	52.5%

Table 2
Detection rates with recovery of missed vertebræ.

Vertebra	Corner type	Detection rate			
		SURF-based		SIFT-based	
C3	Sup.	93.3%	91.3%	76.1%	71.7%
	Inf.	93.3%		71.7%	
C4	Sup.	97.8%	95.7%	84.8%	84.8%
	Inf.	100%		91.3%	
C5	Sup.	100%	95.6%	97.8%	95.6%
	Inf.	97.7%		95.6%	
C6	Sup.	95.4%	93.3%	93.3%	91.1%
	Inf.	97.7%		93.3%	
C7	Sup.	75.0%	72.9%	78.4%	73.0%
	Inf.	75.0%		73.0%	
Mean		92.5%	89.8%	85.5%	83.2%

Table 3
Comparison of the vertebra detection rate with previous methods.

Method	Mode	Detection rate
Dong and Zheng [13]	Semi-automatic	92.4%
Casciaro and Massoptier [14]	Fully automatic	83.0%
Proposed method (SURF-based)	Fully automatic	**89.8%**
Proposed method (SIFT-based)	Fully automatic	83.2%

the size of the descriptor window allowing to obtain the best classification precision. For each vertebral level, we distinguish the corner type: superior and inferior. The percentage indicates the ratio between the number of correctly detected corners by our approach and the total number of corners. The same operation is made for the vertebræ.

In Table 1, the first particular value concerns the C7 vertebra. We notice that, the inferior corner detection rate is very weak. Actually, this is explained by the fact that the C7 vertebra is often partially hidden by other parts of the body. As a result, the anteroinferior corner cannot be easily retrieved during the interest point detection. With regard to the corner detection rates, results are quite satisfying. Nevertheless, the vertebra detection rates are not sufficient (54.6% of success for SURF and 52.5% for SIFT) if we plan to use our framework in a medical context. For this reason, the recovery of missed vertebræ seems essential in our approach.

Table 2 allows to observe the interest of considering an *a priori* on the spine to recover missed vertebræ. The corner detection rates rise from 72.0% to 92.5% for SURF and from 70.1% to 85.5% for SIFT. In a similar way, the vertebra detection rates rise from 54.6% to 89.8% for SURF and from 52.5% to 83.2% for SIFT.

An interesting point is that the SURF-based descriptor is characterized by better detection rates for C3, C4 and C6 vertebræ, while the SIFT-based descriptor presents better results for C7 vertebra. Nevertheless, if we average the detection rates among all the vertebral levels, the SURF-based descriptor shows a mean detection rate greater than the SIFT-based one. As a consequence, we argue that the SURF-based descriptor is to be preferred to the SIFT-based descriptor. The mean classification precision, the mean corner detection rate and the execution time are better for the SURF-based descriptor.

The literature on the subject of fully automatic vertebra detection on X-ray images is poor. We can compare our results to those obtained by Casciaro and Massoptier in [14]. Their approach is fully automatic and the authors showed vertebra detection rate of 83.02%. We notice that our approach shows better values both for SIFT and SURF descriptors. In Table 3, one can observe that only the results obtained by Dong and Zheng in [13] are slightly better. However, their approach is semi-automatic.

4.3. Vertebra segmentation results

In order to assess the precision of our segmentation procedure, we propose to compare the results obtained after a fully automatic initialization of the ASM approach with a semi-automatic initialization. In Table 4, we present the point-to-line distance error (in pixels), the orientation error (in degrees) and the number of failures for the Model 1 and the Model 2. These results have been obtained after a semi-automatic initialization of the ASM approach. A user has been asked to point out the anterior corners of the cervical vertebræ. One can remark that the distance error is given in pixels. Unfortunately, the radiographs we used were not calibrated and it has not been possible to convert the pixels in millimeters. Nevertheless, the information of the distance error is useful to decide which model provides better results.

Similar to the results for vertebra detection, one can notice that the C7 vertebra is characterized by the highest point-to-line distance and orientation errors, regardless of the model. If we compare the results between Model 1 and Model 2, we can remark that Model 2 is always better. Furthermore, this model is more robust since the number of failures for Model 2 is lower than for Model 1. As a consequence, results of Table 4 allows to conclude that Model 2 is to be preferred to Model 1.

Table 4
Segmentation results: semi-automatic initialization.

Vert.	Model 1			Model 2		
	Dist. (px)	Orient. (°)	No. fail.	Dist. (px)	Orient. (°)	No. fail.
C3	4.75	2.46	6	4.63	2.64	3
C4	4.24	2.83	3	4.19	2.42	3
C5	4.51	2.30	3	3.95	1.82	3
C6	3.76	1.60	1	4.40	2.05	1
C7	7.06	4.50	3	5.00	2.53	2
Mean	4.86	2.74	–	4.43	2.29	–

Table 5
Segmentation results: fully automatic initialization.

Vert.	SURF-based			SIFT-based		
	Dist. (px)	Orient. (°)	No. fail.	Dist. (px)	Orient. (°)	No. fail.
C3	5.06	3.30	2	5.03	3.03	1
C4	4.47	2.32	1	4.72	2.91	0
C5	4.21	1.91	0	4.49	2.26	0
C6	4.27	1.89	1	4.67	2.64	4
C7	5.51	3.38	1	5.18	3.18	1
Mean	4.70	2.56	–	4.82	2.80	–

For this reason, the results presented in Table 5 have been obtained thanks to the Model 2. In that case, the initialization of the ASM procedure was based on our fully automatic vertebra detection framework. It is important to clarify that the ASM initialization requires anteroinferior and anterosuperior corners to be located. This condition of initialization is identical for the semi-automatic initialization (Table 4) or for the fully automatic initialization (Table 5). Results of Table 4 actually allow evaluation of the precision of the Model 1 and the Model 2. The results of Table 4 could have been predicted since Model 2 represents the particularities of each vertebral level. This aspect is reinforced for vertebra C7 which is better segmented with Model 2 than Model 1.

The same variables are studied in the case of the fully automatic initialization: the point-to-line distance error (in pixels), the orientation error (in degrees) and the number of failures due to the fully automatic vertebra detection. The two descriptors proposed to initialize the segmentation process are studied: SIFT and SURF. The comments are similar to the previous ones. First, the C7 vertebra is again the one with the worst results. And if we compare the mean point-to-line distance error and the mean orientation error, we note that the SURF-based descriptor is better. Another interesting point is to compare the semi-automatic initialization with the fully automatic one. The point-to-line distance error rises from 4.43 pixels for the semi-automatic approach to 4.70 for the fully automatic based on SURF. Similarly, the orientation error rises from 2.29° to 2.56°. These small differences demonstrate the effectiveness of our fully automatic approach.

A precise landmarking is an important step to establish correspondences between different objects of interest. Here, automated landmarking has been proposed but other possibilities could have been considered. Depending on the subject, three ways of landmarking are possible: manual, semi-automated and fully automated. Manual landmarking should be avoided in our case since it is subjective, time-consuming and prone to errors [28]. Semi-automatic landmarking such as proposed in [29] is an interesting alternative. In that work, an equally spaced landmarking is

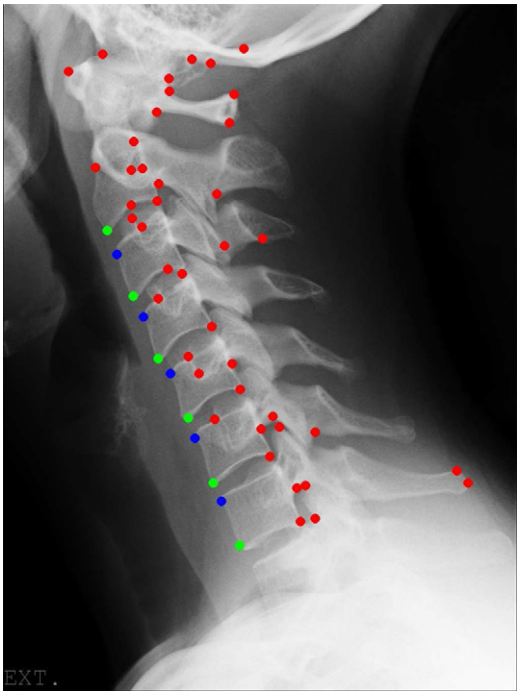

Fig. 4. Vertebra detection: the keypoints are detected and correctly classified. Anterosuperior corners are painted in blue, anteroinferior corners are in green and other detected points are in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 5. Vertebra segmentation: ASM is initialized thanks to vertebra detection. Orange shapes represent the segmented vertebræ. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

operated on shapes based on an initial point given by an operator. Then, a model is constructed so that it encodes the relationship between the shape and the appearance of a medical object. On a new image, the translation, scale and orientation of an anatomical object can be deduced from the intensity structure detected. In our case, the anatomical landmarks would not be equally spaced but could be the corners of the vertebral body. Despite the advantages of the method, images in [29] come from CT or MRI and are focused on a region of interest. In our case, the variability between patients implies that an operator would be required to annotate a region of interest (the cervical spine) on images. On the contrary, our fully automated approach does not need any human intervention on the image. However, a semi-automatic approach could be used as a secondary approach when the fully automatic one fails.

Concerning the efficacy of landmarking, we decided to compare our approach with a manual landmarking operated by a physician. However, the same person does not locate one landmark always in the same place. Furthermore, different persons will locate one landmark in different places. For this reason, we decided to compare our error of landmarking with the inter-operator variability. Success of landmarking is measured from the clinical indices computed from the landmarking. In our case, these clinical indices are the angles of mobility between the cervical segments. Therefore, the efficacy of our fully automatic landmarking will be deduced by the efficacy of computing clinical indices. We will show in Section 4.4 that the angular errors associated to our approach remain within the inter-operator variability range of the reference method.

Finally, we offer a visual illustration of our framework in Figs. 4 and 5. In the former figure, one can see that interest points have been detected on a radiograph. Furthermore, the anterosuperior and anteroinferior vertebra corners have been correctly classified. The latter figure concerns the visualization of the segmentation result on the same radiograph. The segmentation process has been initialized with the vertebra locations determined during the fully automatic vertebra detection.

4.4. Mobility analysis

The framework we propose in this paper is divided into two main steps: vertebra detection and vertebra segmentation. These stages have just been validated in the previous sections. However, our interest is to apply this framework to the analysis of the cervical spine mobility.

In the literature, the most accurate results provide from tests operated on a calibration model and not on real patients. It is the case for [3,5,6,30]. It seems not fair to compare our results with others coming from an assessment on a calibration model and not on real data. In [7], experiments with a videofluoroscopic system on 5 real patients and corresponding angular errors are proposed. Their value reaches 2.52°. Our results for cervical segments are presented in Table 6. Results are logically better for the videofluorosopic system. In our case, the semi-automatic approach is characterized by a mean RMS error of 3.22° while the error for the fully automatic based on SURF is of 3.36°. As a comparison, we observe that our results show slightly greater errors but our approach has the advantage to be fully automatic.

The validation is made by a comparison with a manual marking operated by an experienced radiologist. However, the manual marking is a subjective and tedious step. In [31], the authors showed that the inter-operator variability is about 3.14° for X-ray image

Table 6
Cervical spine mobility analysis: results.

Segment	RMS angle error (°)		
	Semi-auto	SURF-based	SIFT-based
C3/C4	3.10	3.46	3.77
C4/C5	2.97	3.03	4.61
C5/C6	2.92	2.69	2.94
C6/C7	3.90	4.28	4.50
Mean	3.22	3.36	3.95

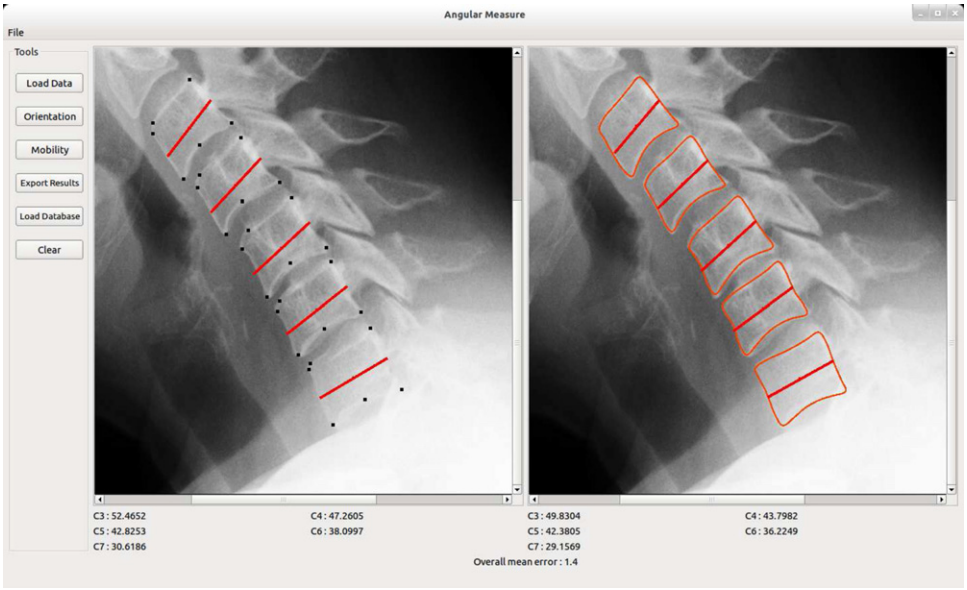

Fig. 6. Cervical spine mobility analysis: graphical interface of our application.

analysis in the context of the vertebral motion analysis. Therefore, the angular errors of Table 6 remain within the inter-operator variability range of the reference method. As a conclusion, we argue that our results, obtained from objective criteria, are acceptable in a medical practice.

We presented in Fig. 6 an illustration of the graphical interface developed in the context of cervical spine mobility analysis. The radiograph on the left has been marked by an experienced radiologist and the vertebra midlines are drawn in red. On the right, vertebræ have been automatically detected and segmented. The same midlines are estimated. The vertebra orientations are shown and compared in the bottom.

5. Conclusion

The objective of this paper was to develop a framework dedicated to the fully automatic cervical spine mobility analysis on X-ray images. To this aim, we proposed an approach based on three main steps: fully automatic vertebra detection, vertebra segmentation and angular measurement.

The vertebra detection is based on a learning method, where the vertebræ are detected by their anterior corners. General interest points are first detected in the radiograph, then are *described* with features such as SIFT or SURF and finally, these points are classified in order to only retrieve anterior vertebra corners. This stage is used in order to initialize the next step: the segmentation procedure based on Active Shape Model. Once the vertebra contours are extracted, angular measurements can be performed in the purpose of studying cervical spine mobility.

Results of vertebra detection are very convincing. We assessed radiographs focused on cervical vertebræ. We have successfully identified the vertebræ in 89.8% of cases. Compared to the literature, these results are very satisfying for a fully automatic method where no human intervention is needed. We have also showed that SURF descriptor slightly outperforms SIFT descriptor in the detection process.

Another interesting conclusion concerns the segmentation results. Our experiments indicated that a statistical shape model specific to the vertebra level is better than a single model for all the vertebræ. Especially for the C7 vertebra which is often partially visible.

Finally, we showed that the results for cervical spine mobility analysis are acceptable for a medical practice. The interest is that they are obtained from objective criteria and in a fully automatic way. We showed that angular errors remain within the inter-operator variability of a manual marking.

In the future, it could be relevant to transpose this framework to a digital videofluoroscopy system. Since the image quality is different between conventional radiographs and DVF images, the methods should be adapted accordingly.

Conflict of interest

None.

Acknowledgements

We would like to thank Paul Desclée, Physician in the Radiology Department of the Jolimont Hospital, for having provided and annotated the radiographs. The validation of our approach would not have been possible without his help. We are also grateful to Aloys du Bois d'Aische for insightful discussions on this work. Finally, we would like to thank the anonymous reviewers for their valuable comments helping to improve the paper.

References

[1] Cerciello T, Bifulco P, Cesarelli M, Romano M, Allen R. Automatic vertebra tracking through dynamic fluoroscopic sequence by smooth derivative template matching. In: 9th international conference on information technology and applications in biomedicine. 2009.
[2] Breen AC, Allen R, Morris A. Spine kinematics: a digital videofluoroscopic technique. J Biomed Eng 1989;11(3):224–8.
[3] Bifulco P, Cesarelli M, Allen R, Sansone M, Bracale M. Automatic recognition of vertebral landmarks in fluoroscopic sequences for analysis of intervertebral kinematics. Med Biol Eng Comput 2001;39:65–75.
[4] Muggleton JM, Allen R. Automatic location of vertebrae in digitized videofluoroscopic images of the lumbar spine. Med Eng Phys 1997;19(1):77–89.
[5] Cerciello T, Romano M, Bifulco P, Cesarelli M, Allen R. Advanced template matching method for estimation of intervertebral kinematics of lumbar spine. Med Eng Phys 2011;33:1293–302.
[6] Zheng Y, Nixon MS, Allen R. Automated segmentation of lumbar vertebrae in digital videofluoroscopic images. IEEE Trans Med Imaging 2004;23(1):45–52.
[7] Lam SCB, McCane B, Allen R. Automated tracking in digitized videofluoroscopy sequences for spine kinematic analysis. Image Vision Comput 2009;27:1555–71.

[8] Reinartz R, Platel B, Boselie T, van Mameren H, van Santbrink H, Romeny BH. Cervical vertebrae tracking in video-fluoroscopy using the normalized gradient field. Med Image Comput Comput Assist Interv 2009;12: 524–31.
[9] Frobin W, Leivseth G, Biggemann M, Brinckmann P. Sagittal plane segmental motion of the cervical spine. A new precision measurement protocol and normal motion data of healthy adults. Clin Biomech 2002;17: 21–31.
[10] Puglisi F, Ridi R, Cecchi F, Bonelli A, Ferrari R. Segmental vertebral motion in the assessment of neck range of motion in whiplash patients. Int J Legal Med 2004;118(4):235–9.
[11] Benjelloun M, Mahmoudi S. X-ray image segmentation for vertebral mobility analysis. Int J Comput Assist Radiol Surg 2008;2(6):371–83.
[12] Zamora G, Sari-Sarraf H, Long LR. Hierarchical segmentation of vertebrae from X-ray images. In: Proceedings of the SPIE medical imaging, vol. 5032. 2003. p. 631–42.
[13] Dong X, Zheng G. Automated vertebra identification from X-ray images. In: Image analysis and recognition, Lecture notes in computer science, vol. 6112. 2010. p. 1–9.
[14] Casciaro S, Massoptier L. Automatic vertebral morphometry assessment. In: 28th annual international conference of the IEEE engineering in medicine & biology society. 2007. p. 5571–4.
[15] Lecron F, Benjelloun M, Mahmoudi S. Fully automatic vertebra detection in X-ray images based on multi-class SVM. In: Proceedings of the SPIE medical imaging, vol. 8314. 2012.
[16] Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models-their training and application. Comput Vision Image Underst 1995;61(1):38–59.
[17] Lecron F, Benjelloun M, Mahmoudi S. Points of interest detection in cervical spine radiographs by polygonal approximation. In: International conference on image processing theory, tools and applications. 2010. p. 81–6.
[18] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vision 2004;60(2):91–110.
[19] Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). Comput Vision Image Underst 2008;110(3):346–59.
[20] Vapnik VN. The nature of statistical learning theory. New York: Springer-Verlag; 1995.
[21] Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE Trans Pattern Anal Mach Intell 2001;23(6):681–5.
[22] Benjelloun M, Mahmoudi S, Lecron F. A framework of vertebra segmentation using the active shape model-based approach. Int J Biomed Imaging 2011: 1–14.
[23] Gururajan A, Kamalakannan S, Sari-Sarraf H, Shahriar M, Long R, Antani S. On the creation of a segmentation library for digitized cervical and lumbar spine radiographs. Comput Med Imaging Graph 2011;35(4):251–65.
[24] Iglesias JE, de Bruijne M. Semiautomatic segmentation of vertebrae in lateral X-rays using a conditional shape model. Acad Radiol 2007;14(10):1156–65.
[25] Smyth PP, Taylor CJ, Adams JE. Vertebral shape: automatic measurement with active shape models. Radiology 1999;211(2):571–8.
[26] Roberts MG, Cootes TF, Adams JE. Vertebral morphometry: semiautomatic determination of detailed shape from dual-energy X-ray absorptiometry images using active appearance models. Invest Radiol 2006;41:849–59.
[27] Roberts MG, Oh T, Pacheco E, Mohankumar R, Cootes TF, Adams JE. Semiautomatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. Osteoporos Int 2012;23:655–64.
[28] Davies R, Twining C, Taylor C. Statistical models of shape: optimisation and evaluation. 1st ed. London: Springer; 2008.
[29] Bagci U, Chen X, Udupa JK. Hierarchical scale-based multiobject recognition of 3-d anatomical structures. IEEE Trans Med Imaging 2012;31(3):777–89.
[30] Breen AC, Muggleton JM, Mellor FE. An objective spinal motion imaging assessment (OSMIA): reliability, accuracy and exposure data. BMC Musculoskelet Disord 2006;7(1).
[31] Breen AC, Teyhen D, Mellor F, Breen A, Wong K, Deitz A. Measurement of intervertebral motion using quantitative fluoroscopy: report of an international forum and proposal for use in the assessment of degenerative disc disease in the lumbar spine. Adv Orthop 2012:1–10.

Fabian Lecron received the computer science engineering degree from the Faculté Polytechnique de Mons, Belgium, and the management science degree from the Facultés Universitaires Catholiques de Mons (FUCaM), Belgium, respectively in 2008 and 2011. He is now preparing a PhD thesis in the field of statistical modeling of the spine applied to medical image analysis at the University of Mons (formerly Faculté Polytechnique de Mons), Belgium. His main research areas are computer vision, image processing and pattern recognition

Mohammed Benjelloun received the electromechanic engineering degree and the management and computer engineering degree respectively in 1986 and 1989. He completed his PhD in applied sciences in 1994 from the Faculty of Engineering, Mons, Belgium (FPMs). He is currently associate professor in the Department of Computer Science at the Faculty of Engineering FPMs of the University of Mons. His research interests include computer vision, image processing, 2D and 3D retrieval, parallel computing and software engineering. He was involved in several projects including industrial and academic researchers and he participates in numerous scientific and academic activities.

Saïd Mahmoudi received in 1997 the B.S. degree in computer science engineering from the University of Science and Technology of Oran (SENIA), Algeria, and in 1999 the M.S. degree in computer science from the LIFL Laboratory, University of Lille 1, France. After that, he joined the FOX-MIIRE Group of the LIFL Laboratory, at the University of Science and Technology of Lille 1, France, where he started his Ph.D. thesis. During this period, he worked on image recognition and 3D retrieval using characteristic views. In 2003, Dr. Mahmoudi obtained the Ph.D. degree in the Computer Science at the University of Science and Technology of Lille 1, France. Between 2003 and 2005 he was associate lecturer at the University of Lille 3, France. Since September 2005, he is associate professor at the Faculty of Engineering of the University of Mons, Belgium. His research interests include medical images processing and computer aided diagnosis, 2D and 3D retrieval and indexing.