Strings and necklaces: on learning and browsing medical image segmentations
Ghebreab, S.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Medical Image Retrieval by Browsing Population-Based Incrementally Learned String Segmentations*

This chapter proposes a browsing method for medical image retrieval by example objects. The characteristics of a population of normal objects are inductively learned by functional data analysis and summarized in a string segmentation model. The string model is used for segmentation of an example image of an abnormal object at the beginning of the browsing session. The segmentation result is used to adapt the string model towards one that can handle the recorded abnormality. The segmentation result, in addition, bootstraps retrieval of images containing similar abnormal objects. In an iterative process, good or specifically bad retrievals are segmented according to the new string model in order to refine the definition of the abnormality. The segmentations form a pilot for a population-based incremental learning technique that explores the repository and exploits previous retrieval results to arrive at a population of objects with the same common deviations from normal. Browsing ends with a set of images depicting same abnormalities as the initial example, ranked by the degree of content matching. The method has been successfully applied for retrieval of digitized X-ray images of vertebrae from a large collection of normal and abnormal instances.

*Submitted to IEEE Transactions on Medical Imaging
4.1 Introduction

Medical images come from an ever-increasing variety of sources and in ever-increasing amounts, bootstrapping the development of new techniques for content-based image retrieval. Traditionally, retrieval techniques utilize textual information in the patient record to recover images from a repository. This approach has proven to be successful for medical imaging [75], but it is apparent that the rich information contained in the images themselves is not fully exploited. Medical image data sets demand a higher degree of content understanding [118] to fully bestow visual information. In this context, content-based image retrieval by example provides an opportunity to tap the expertise present in an image repository based on visual features, once the textual information in the patient record is exhausted. We take retrieval one step further by explaining images based on objects contained in them. In its impact this is radically different from mainstream content-based image retrieval [111] where information is summarized from the entire image disregarding object properties.

A number of methods in literature, e.g. [25],[30], [57], [72], [76], [91], [103], [104], [106], try to tackle the retrieval problem in an object-based manner using geometrical, analytical or grammatical segmentation models. The problem with these approaches is the premise that the segmentation model is capable of capturing an arbitrary object in an image, regardless of its natural variations. Apart from this, the approaches postulate that a single example object suffices to retrieve relevant images from the repository. This might be the case when objects have clear and consistent characteristics, easily distinguishing them from other ones. However, if this assumption is violated, as is often the case in medical imaging, segmentation models will fail due to their incapacity to precisely capture the user’s retrieval intention, by learning from multiple query objects, user-interaction or from the history of a browsing session. Learning the search intention of the user is an important problem with direct implications on the ability of an algorithm to meet the user’s demand.

The problem addressed in this work is: how to learn the user’s retrieval intention and recover images best matching his/her concept from large collections of images, commencing from a single example object? The approach taken departs from the view that a) to accurately capture objects in images, segmentation models should be learned in terms of multiple features rather than constructed from a priori knowledge, b) to handle arbitrary objects, segmentation models should be adaptive rather than fixed c) to readily circumscribe the user’s concept of the object, it should be described with help of multiple example objects exhibiting similar characteristics. In this context, medical image retrieval requires learning the characteristics of normal and abnormal object populations. The problem of retrieval is transposed into one of browsing the image repository in search for abnormal populations with the same deviating characteristics from normal as the example object.

The chapter is organized as follows. Section 2 describes the image material and the application to demonstrate the proposed retrieval method. Section 3 is devoted to the proposed retrieval scheme. The main issues are inductive learning, string-based segmentation, population-based incremental learning and content-based image retrieval. In section 4 automatic image classification is addressed. Experiments and
results are described in section 5, then discussion and conclusion in section 6.

4.2 Clinical Example Application

The application to demonstrate the retrieval problem and proposed solution deals with cervical vertebra images exhibiting osteophytes. The image material is acquired from the approximately 17,000 X-ray films collected during the Second National Health and Nutrition Examination Survey (NHANES II) conducted by the NCHS [86]. In this cross-sectional population survey, X-rays were taken of persons aged between 25 and 74. Two X-rays of the spine, PA and lateral, were made except of pregnant women and women under 50 years of age, to provide evidence of osteophyte and degenerative disc diseases. The films were subsequently digitized at a horizontal and vertical sampling rate of 146 dpi using Lumisys laser scanning equipment [86].

![Digital images of cervical vertebrae](image_url)

**Figure 4.1:** Digital images of cervical vertebrae: a) normal, b) lower osteophyte, c) upper osteophyte, d) lower and upper osteophyte. Note the limited detail and the complexity of the boundary.

We concentrate on a repository consisting of 283 digitized NHANHES images of single cervical vertebrae. A medically certified diagnosis is attached to each image, including an expert delineation of the vertebra. The image repository contains both normal and abnormal vertebrae. The abnormal vertebrae are classified either as lower anterior osteophyte, upper anterior osteophyte or both lower and upper anterior osteophyte. To illustrate the differences between them consider figure 4.1.

An osteophyte is characterized by bony outgrowths on the anterior corners of the vertebral body. For example, the shape of the lower anterior of the C5 vertebra in figure 4.1b clearly extends from the body of the vertebra. The spurs are furthermore associated with the structural image boundary as opposed to histogram or texture characteristics [97]. We utilize shape and edge information related to the vertebra boundary for describing and retrieving images by visual features.

We simulate a situation in which a clinician wants to retrieve vertebra images
from the repository of known and annotated cases to help determine the diagnosis of an unknown abnormal vertebra image, in this case one classified as lower anterior osteophyte. The clinician uses the abnormal vertebra as an initial example for retrieval similar images from the repository. The clinician then refines the initial query by indicating good and bad retrievals as new examples for the next retrieval step, this way browsing through the retrieved images. This brings the clinician to a population of images best matching the initial retrieval intention, hence best suited for supporting the diagnosis. This is a highly challenging task considering the subtle shape and image differences among vertebrae and in this case the marginal image quality.

### 4.3 Method for Browsing Image Populations

The proposed image retrieval scheme consists of four components as illustrated in figure 4.2. *Inductive learning* constructs a string model of a normal vertebra by learning from a given set of images with true segmentations. The model is used for *string-based segmentation* of the example image of the unknown abnormal vertebra at the beginning of the browsing session. The segmentation result is used to adapt the string model towards one that can handle the recorded abnormality. The segmentation result, in addition, is used for *content-based retrieval* of images depicting similar abnormalities. To refine the definition of the abnormality and hence retrieval, good or bad retrievals are indicated as new search examples. The examples form a pilot for the *population-based incremental learning* technique used to explore the repository and to exploit previous results. Browsing ends with a set of images depicting same abnormalities as the initial query image, ranked by the degree of content matching.

In the following sections we address each of the components in figure 4.2. We start with a brief description of the string segmentation model proposed in the previous chapter, which in contrast to other similar approaches (e.g. [11], [19], [20], [35], [43], [85]) has the capacity to use multiple continuous shape and image features in the definition of boundaries.

#### 4.3.1 Defining a Population of Normals

We analyze a population of normal vertebrae in order to learn the important statistical characteristics of this population. We do this inductive learning on the basis of multiple continuous features recorded along normal vertebra boundaries.

Vertebra boundaries in the training images are represented by smooth curves $s : \mathbb{R} \rightarrow \mathbb{R}^2$ parameterized by $v \in \mathbb{R}$. Given the population of $M$ input images $I_1(x), \ldots, I_M(x)$ for each of which the true shape $s_1(v), \ldots, s_M(v)$ is known, the learning set consists of pairs of image and shape data

$$\mathcal{L} = \{(I_1(x), s_1(v)), \ldots, (I_M(x), s_M(v))\}. \tag{4.1}$$

For the $m$th learning example, the shape $s_m(v)$ relates to the image at vertebra boundary points $I_m(s_m(v))$. The relation is expressed in terms of $N$ features measured
Figure 4.2: Overview of the retrieval scheme, consisting of components for inductive learning, string-based image segmentation, incremental learning and content-based image retrieval.

along $I_m(s_m(v))$ derived from the shape as well as from the image. This yields $M$ feature functions in the $N$-dimensional feature space $\mathcal{F}$

$$f^{**}_m(v) = [f^{**}_{m1}(v), \ldots, f^{**}_{mN}(v)].$$ (4.2)

The feature functions are aligned to remove variation from the learning data which is attributed to stretching, shearing and rotation of the feature data. This is accomplished using the Procrustes method [50] by finding for $f^{**}_m(v)$ the strictly monotonic non-linear transformation $\omega_m(v)$ [95] that produces the warped feature function

$$f^*_m(v) = f^{**}_m(\omega_m(v)).$$ (4.3)

The warping functions and the aligned feature functions are computed iteratively with help of the estimated mean feature function $\hat{\mathbf{f}}(v)$. A global alignment criteria is used for computing $\omega_m(v)$ for $f^{**}_m(v)$ such that

$$\omega_m(v) = \arg\min_{\omega^*_m(v)} \sum_{m=1}^{M} \int_v ||f^{**}_m(\omega^*_m(v)) - \hat{\mathbf{f}}(v)||^2 dv. \quad (4.4)$$

The final mean feature function $\hat{\mathbf{f}}(v)$ is computed from this aligned set of feature functions and subsequently subtracted from each function $f^*_m(v)$ to obtain feature functions $f_m(v)$, containing all evidence of the learning phase properly aligned, cen-
tered and scaled to unit variance. The normalized feature data is defined by

$$f_m(v) = \frac{f_m^*(v) - \bar{f}(v)}{\sigma_f(v)}$$

(4.5)

with units of variance due to normalization by

$$\sigma_f(v) = \left( \frac{1}{M} \sum_{m=1}^{M} ||f_m^*(v) - \bar{f}(v)||^2 \right)^{1/2}$$

(4.6)

Principal components analysis is performed to project the $N$-dimensional functional data to a smaller feature space. Assuming the number of main modes of variation is $Q$, weighting functions $\alpha_q(v) = [\alpha_{q1}(v), ..., \alpha_{qN}(v)]$ are sought for one-by-one, to find values

$$g_{mq} = \sum_{n=1}^{N} \int \nu f_{mn}(v) \alpha_{qn}(v) dv$$

(4.7)

that explain most of the variation in the data. The function $\alpha_{qn}(v), q = 1, ..., Q$ is a weight chosen so as to highlight variation in the data in dimension $n$. The values $g_{mq}$ form the principal component score $g_m = [g_{m1}, ..., g_{mQ}]$, used in place of the actual feature function $f_m(v)$ to summarize the bulk of feature data to a vector-valued quantity. The vector of weighting functions $\alpha_q(v) = [\alpha_{q1}(v), ..., \alpha_{qN}(v)]$ is computed such that

$$\alpha_q(v) = \operatorname{argmax}_{\alpha_q^*(v)} \frac{1}{M} \sum_{m=1}^{M} g_{mq}^2$$

(4.8)

Principal component regression [14] is performed to build an underlying model of the data in the learning set. The matrix of regression functions $B(v) = [\beta_1(v), ..., \beta_Q(v)]^T$ is computed by least squares minimization such that

$$B(v) = \operatorname{argmin}_{B^*(v)} \sum_{m=1}^{M} \int \nu ||f_m(v) - g_m B^*(v)||^2 dv.$$  

(4.9)

The matrix $B(v)$ indicates how each of the $N$ measured features along the vertebra contour contributes to the principal component functions. Hence, the regression functions indicate which features are locally most important to define the vertebra characteristics. With help of the estimated regression functions, principal component scores are predicted for unknown feature function emanating from an unknown vertebrae.

A Mahalanobis distance model is computed from the $Q \times M$ matrix $G$, composed of the $q$-dimensional principal component scores, one for each $M$ feature functions. This differs from chapter 3 in that we omit residual information derived from minimizing equation 4.9. We omit this information in order to allow comparison of virtual feature
functions with no residuals (yet to be described), with real feature functions that produce a residual when explained by the regression model. The Mahalanobis distance matrix is then defined as

$$D = \frac{G^T G}{(M + 1)}.$$  \hspace{1cm} (4.10)

The Mahalanobis distance matrix allows to compute the distance of the individual $f_m(v)$ or any other feature function to the population average, taking into account variation. In this sense, the Mahalanobis distance is a measure of quality of an individual with respect to an entire population. For $f_m(v)$ the Mahalanobis distance is computed using $g_m$

$$D^2(f_m(v)|g_m) = g_m D^{-1} g_m^T.$$ \hspace{1cm} (4.11)

To make the Mahalanobis distance computed in equation 4.11 only dependent on the shape of a cluster, not its size, we remove the scale difference using the root mean squared group size \cite{32}, contrary to what is done in chapter 3. To obtain the scale invariant Mahalanobis distance the quantity in 4.11 is divided by

$$R = \frac{1}{M - 1} \sum_{m=1}^{M} D^2(f_m(v), \bar{f}(v)).$$ \hspace{1cm} (4.12)

The following information is kept from the inductive learning phase: the population average $\bar{f}(v)$, the normalization function $\sigma_f(v)$, the matrix of regression functions $B(v)$ and the principal components scores matrix $G$. This information is transferred to the segmentation phase for finding a vertebra boundary in an unknown image on the basis of multiple weighted features.

### 4.3.2 Visualizing the Normal Population

We illustrate and discuss the appearance of a normal cervical vertebra according to the string model. In order to get better insight in the shape of the normal cervical vertebra and the natural variations therein, we first concentrate on $N = 2$ shape features, capturing the projectional alignment of vertebrae in addition to local shape.

We visualize the learning data in the image space. The images in the top row of figure 4.3 illustrate the average shape of a normal cervical vertebra plus (light grey values) and minus (dark gray values) up to three standard deviations away from the average shape. The four sub-figures 4.3a-4.3d correspond to variation in the first four principal component directions, together capturing 93 percent of the total variability in shape. The number of principal components has been set to $Q = 4$ because we expect for our vertebra application there are 4 corners, and hence 4 places, in the string model where the data in the learning phase might exhibit independent shape variation.

Figure 4.3a shows that the main variation in the shape of the normal cervical vertebrae in our application occurs at the lower anterior corner. The variation extends
Figure 4.3: Visualization of shape and image characteristics of a population of normal cervical vertebrae. Top row visualizes the average shape plus (light) and minus (dark) up to three standard deviation in the a) first, b) second c) third and d) fourth principal components direction. Bottom row visualizes the average gradient along the average vertebra contour plus (positive normal) and minus (negative normal) up to three standard deviations away in the of e) first f) second g) third and h) fourth principal components direction.

For the gray-level definition of the cervical vertebra boundary we consider $N = 1$ image feature. The image gradient magnitude along the vertebra boundary is recorded. The bottom row in figure 4.3 visualizes along the average vertebra contour the average gradient magnitude value and the variation therein. The variation is visualized in the normal direction of the average contour by plotting the average gradient value plus a multiple of a principal component in the positive normal direction.
and minus a multiple in the negative normal direction. The brighter the intensity the better defined the image gradient. The broader the scope of the intensity in perpendicular direction the more variation there is in gradient magnitude.

Figure 4.3e shows that the image gradient is well-defined at the upper and lower anterior corner and upper right corner of the vertebral body. As expected the gradient magnitude is ill-defined due to absence of intensity discontinuity at the part where the vertebral body joins the pedicles. The image gradient is also less steep at the middle upper and lower parts. It is not immediately visible where the main modes of variation occur. However, upon inspection of the image differences (data not shown here) it becomes clear that the first mode of variation focuses on the upper right corner, the second on the lower anterior corner, the third on the entire lower part of the vertebral body and the fourth along the entire right part of the vertebral body. From these observations we conclude that there is some correlation between shape and diagnostically significant image features.

4.3.3 String-Based Image Segmentation

At the beginning of a browsing session we need to segment the abnormal vertebrae in the unknown example image in order capture the characteristics of that vertebra. String-based image segmentation allows to do this in terms of multiple features, weighted according to the most important variations in the normal population.

An active string in the $N$-dimensional feature space $\mathcal{F}$, denoted by $f_t^{**}(v)$, is used for image segmentation and feature extraction. This string is defined by features extracted from a shape model $s_t(v)$ that deforms in time $t$ and from the unknown vertebra image $I_0(x)$ in which that model lives. At time $t$ the feature function is defined as

$$f_t^{**}(v) = [f_{11}^{**}(v|s_t(v), I_0(x)), ..., f_{N}^{**}(v|s_t(v), I_0(x))].$$  \hfill (4.13)

After aligning $f_t^{**}(v)$ to the population average $\bar{f}(v)$ and weighting it by $\tilde{f}(v)$, both derived from the learning phase, its quality with respect to the learning feature functions $f_1(v), ..., f_M(v)$ is determined from the relation of its corresponding score vector $g_t$ to the cluster of principal component scores $g_1, ..., g_M$. First, the $Q$-vector $g_t$ is estimated by solving

$$f_t(v) = g_tB(v)^T + \epsilon_t(v)$$  \hfill (4.14)

where $\epsilon_t(v)$ is the residual function. This allows estimation of $g_t$ such that the important features of $f_t(v)$ according to the regression model in equation 4.9 are emphasized. The value of $g_t$ is estimated by least squares minimization such that

$$g_t = \arg\min_{g_t} \int_v (g_tB(v)^T - f_t(v))^2 dv.$$  \hfill (4.15)

The Mahalanobis distance of $f_t(v)$ to the average of the normal population is then defined in terms of the distance of the new score $g_t$ to the average of the cluster
constituting of \( g_1, ..., g_M \). It is computed as follows

\[
D^2_M(f_t(v), \tilde{f}_t(v)) = g_t D^{-1} g_t^T. \tag{4.16}
\]

We use this quality measure for segmenting the unknown image \( I_0(x) \). The string \( f_t(v) \) vibrates in feature space due to deformations of the shape model \( s_t(v) \) in the image, subsequently producing new scores evaluated by means of the Mahalanobis distance model. Starting from an initial configuration, \( s_t=0(v) \) is deformed by tuning its shape parameters in such a way that the state of minimal energy provides the optimal shape model \( s_f(v) \), assumed to be the outline of the vertebrae in the unknown image. This reduces to optimizing \( f_t(v) \), emanating from \( I_0(s_t(v)) \), such that

\[
f(v) = \arg\min_{f_t(v)} D^2_M(f_t(v), \tilde{f}(v)). \tag{4.17}
\]

At this point we have: a string model trained from a population of normal vertebra images and a string segmentation of an unknown abnormal vertebra for which we try to find similar images in the repository of known cases. The string segmentation of the abnormal vertebra is used to bootstrap the browsing image segmentations. The string model is refined during browsing in order to more precisely define the characteristics of the browsed population, i.e. of the abnormality.

### 4.3.4 Browsing a Population of Abnormals

As the string segmentation of the abnormal vertebra in the initial example image may not suffice to unequivocally describe the characteristics of that vertebra, the image repository of known cases is browsed for similar images to learn those characteristics more carefully. We do this by incrementally learning the vertebra features recorded in the retrieved images that are interactively specified as good or as bad retrievals.

For browsing we use a population-based probabilistic search algorithm based on modeling promising areas in the feature space by estimating their probabilistic distribution [3], [4]. The probabilistic model explores the feature space taking into account information from previous explorations. At the end of the exploration, the probabilistic model defines a narrow region of the feature space, frequently encountered during the exploration process, hence circumscribing the retrieval intention in terms of reoccurring characteristics.

We conceive of the feature set used for inductive learning as an initial population, bounding a feature subspace that defines the appearance of a normal population. That is, feature functions \( f^{**}_1(v), ..., f^{**}_M(v) \) constitute the first generation of an evolving population. For simplicity of notation, from this point on we denote the evolving population by \( f^s_t(v), ..., f^s_M(v) \). At generation \( s = 0 \) we have

\[
f^s_m(v) = f^{**}_m(v) \tag{4.18}
\]

with the corresponding principal component scores denoted by

\[
g^s_m = g_m. \tag{4.19}
\]
4.3. Method for Browsing Image Populations

The probabilistic model \( P^s(f^*_m | g^*_m) \) corresponding with this population is defined using Gaussian models with mean \( \bar{f}^*(v) \) and standard deviation \( \sigma^*_f(v) \). This is permissible as we expect a Gaussian distribution of feature functions. The probability of \( f^*_m(v) \) is computed on the basis of its principal component scores as follows

\[
P^s(f^*_m | g^*_m) = |D|^{-1/2} (2\pi)^{-Q/2} e^{-D^{-1} s_m^T s_m}. \tag{4.20}
\]

With help of \( P^s(f^*_m | g^*_m) \) we maintain a probabilistic model of the promising regions of the search space. The aim is to adapt \( P^s(f^*_m | g^*_m) \) by exploration and exploitation of the feature space such that at the end of the evolution, vertebra with abnormalities similar to the initial and subsequent examples have high probability according to this model. Representing the \( u \)th example vertebra by feature function \( f^*_u(v) \) for \( u = 1, ..., U \) and corresponding score by \( g^*_u \), this means we want to find the probabilistic model \( P^*(f^*_u | g^*_u) \) with

\[
P^*(f^*_u | g^*_u) = \arg\max_{P^*(f^*_u | g^*_u)} \sum_{u=1}^U P^s(f^*_u | g^*_u). \tag{4.21}
\]

The \( U \) example vertebrae \( f^*_u(v) \) form the pilot for evolution, with the \( u \)th query example corresponding with the \( s \)th browsing step. In the initial browsing step the query example is the abnormal vertebrae in the unknown image. That is, for \( s = 1 \) and hence for \( u = 1 \), the query example is defined by the score \( g^*_u \) obtained by optimization of a string in the unknown image \( I_u(x) \) such that

\[
f^*_u(v) = \arg\min_{f^*_u(v)} D^2_M(f^*_u(v), \bar{f}(v)). \tag{4.22}
\]

To determine whether a feature subspace corresponds to the concept of the sought abnormal vertebra we define an evaluation function based on the most promising individual of the population at hand. In the initial browsing step at \( s = 1 \) the promising individual is the one most resembling the abnormal vertebrae in the initial query image. The score \( g^*_u \) is used as a reference for selecting this individual. To this end a quantity is defined that measures the Euclidean distance between an individual \( g^*_m \) and \( g^*_u \)

\[
D_E(f^*_m(v), f^*_u(v) | g^*_m, g^*_u) = \|g^*_m - g^*_u\|. \tag{4.23}
\]

In an elitist approach, the individual is selected that has smallest Euclidean distance to \( g^*_u \). This individual forms the center of a new feature subspace that presumably better describes the features of the abnormal vertebra in the initial query than does \( f^*(v) \). The new feature function is selected to define the new mean \( \bar{f}^*(v) \) for the new probabilistic model

\[
\bar{f}^*(v) = \arg\min_{1 \leq m \leq M} D_E(f^*_m(v), f^*_u(v)). \tag{4.24}
\]

We also employ the least matching individual of the population so that we can rely on negative examples in addition to a search based on positive examples. This allows to
explicitly formulate that the search is not targeted at normal vertebrae or at vertebrae with abnormalities not related to the initial example. Defining the new mean on the basis of the worst individual reduces to

\[
\bar{f}^*(v) = \arg\max_{1 \leq m^* \leq M} D_E(f^*_m(v), f^*_u(v)).
\]  

As the aim is definition of the new probabilistic model for the next generation we also need \(\sigma^*_f(v)\) in addition to \(\bar{f}^*(v)\). In the initial browsing step \(\sigma^*_f(v)\) is initialized to some large value in an attempt to ensure that the feature space is well-covered in terms of probability. In the subsequent browsing steps it is annealed towards a value that represents the variation in the browsed population. That is, given examples \(f^*_u(v)\) for \(u = 1, \ldots, U\), we have

\[
\sigma^*_f(v) = \left( \frac{1}{U} \sum_{u=0}^U \| f^*_u(v) - \bar{f}^*(v) \|^2 \right)^{1/2}
\]  

(4.26)

The probabilistic model \(P^*(f^*_m|g^*_m)\) corresponding to the newly suggested feature subspace is defined by \(\bar{f}^*(v)\) and \(\sigma^*_f(v)\). The probability model \(P^{s+1}(f^*_{m+1}|g^*_{m+1})\) for the next generation is updated using information gained from \(P^s(f^*_m|g^*_m)\) and \(P^*(f^*_m|g^*_m)\). The probability update rule employed here is similar to weight update rule in competitive learning when an output is moved towards a particular sample feature function \([3]\). The next generation probability model is defined by mean

\[
f^{s+1}(v) = (1 - \gamma) \bar{f}^s(v) + \gamma \bar{f}^*(v)
\]  

(4.27)

and deviation

\[
\sigma^{s+1}_f(v) = (1 - \gamma) \sigma^s_f(v) + \gamma \sigma^*_f(v)
\]  

(4.28)

The innovation rate \(\gamma\) offers explicit control of how fast the population converges. As the probability model is used to generate the next population, the learning rate also affects which portion of the feature space will be explored. It updates the probability model in the direction of the best individual of the current population, governed by \(\gamma\). When the learning rate is 0, there is no exploitation of the information gained through search. As the learning rate is increased, the amount of exploitation increases, and the ability to search large portions of the feature space diminishes.

The next population of feature functions arises from the updated probability model \(P^{s+1}(f^*_{m+1}|g^*_{m+1})\). The feature functions are not real in the sense that they do not necessarily relate to real vertebrae. Rather they are virtual functions arising from exploring the feature space, possibly corresponding with unrealistic vertebrae. The virtual feature functions are obtained by sampling \(M\) times from \(P^{s+1}(f^*_{m+1}|g^*_{m+1})\)

\[
f^{s+1}_m(v) \sim P^{s+1}(f^{s+1}_m, \sigma^{s+1}_f(v)).
\]  

(4.29)

To continue exploring the feature space by evolving the probabilistic model, we need further query examples that are similar or specifically dissimilar to the abnormal vertebra in the initial query image. We fetch the example from the image repository.
4.3. Method for Browsing Image Populations

To this end, feature functions $f_b(v)$ corresponding to the $b = 1 \ldots B$ vertebra images in the repository are evaluated using a string model that is trained from all $M$ virtual feature functions $f_{m+1}(v)$ by performing the same steps as described in equations 4.3-4.12. The principal component score $g_{b}^{s+1}$ corresponding with the $b$th repository instance, computed according to the string model describing the population at generation $s + 1$, is obtained by solving the following equation

$$f_b(v) = g_b^{s+1} B^{s+1}(v)^T + \epsilon^{s+1}(v)$$

using least squares minimization of the residual $\epsilon^{s+1}(v)$. In here $B^{s+1}(v)$ is the matrix of the re-computed regression functions. This way feature functions $f_b(v)$ are reweighted after each browsing step so as to qualify them according to the common features recorded in the newly obtained population. To determine the promising images in the repository, we compute the Mahalanobis distance of each score $g_b^{s+1}$
to the average of the evolved population on the basis of the updated distance model $D^{s+1}$

$$D_M^2(f_b(v), \bar{f}^{s+1}(v)) = g_b^{s+1} D^{s+1-1} g_b^{s+1T}. \quad (4.31)$$

The images in the repository with minimal Mahalanobis distance are the ones similar to the vertebra in the initial query image. For the next browsing step the images in the repository are ranked according to their Mahalanobis distance so that the most similar are placed first and the least one at the last place. The prime image suggested for browsing is the one with minimal Mahalanobis distance

$$f_b(v) = \arg\min_{1 \leq b^* \leq B} D_M^2(f_b(v), \bar{f}^{s+1}(v))/R^{s+1}. \quad (4.32)$$

The above process of incrementally updating the probability model and the string model, and hence redefining and refining the concept of the sought abnormal vertebra, is repeated a number of times, until browsing has narrowed the feature space such that a sufficient number of relevant images of abnormal vertebrae from the image repository are fetched. A diagnosis can then be attached to the unknown abnormal vertebra in the initial example image based on the known classifications of the retrieved images.

### 4.3.5 Visualizing the Browsed Population

To demonstrate how the vertebra shape evolves when browsing the image repository with images of lower anterior osteophyte as the target, we visualize the vertebrae shapes emanating from the virtual feature functions $f_m^{s+1}(v)$. We have selected an image of a vertebra possessing a lower anterior osteophyte which we know to be present in the image repository. Hence, the string model of the normal cervical vertebra is evolved to a model of a vertebra with a lower anterior osteophyte.

The shapes in the top row of figure 4.5 indicate the evolution trajectory from normal cervical vertebra to one which possesses a lower anterior osteophyte, requiring only four browsing steps. From the figures it can be seen which shapes are gradually explored to arrive at lower anterior osteophyte, with the modes of variation showing the parts of the vertebra contour that are adapted during browsing. It can be seen from figure 4.5d that there is almost no variation at the end of the browsing session as, in this case, the same abnormal vertebra has been used for exploration. When exploring the image repository on the basis of a population of abnormalities, the average indicates the common characteristics of the browsed normals, whereas the modes of variation indicate parts of the abnormal vertebrae where the characteristics are equivocal. It might be postulated that the evolution from normal to abnormal vertebra perhaps corresponds to some pathological process.

The example demonstrates that the population-based incremental learning algorithm fetches relevant images from the repository within a small number of browsing steps. In this case the innovation rate $\gamma$ for the trade-off between exploration and exploitation has been set to emphasis exploration rather than exploitation, allowing to browse quickly from the population of normal vertebrae to the population with lower anterior osteophyte. When the target of browsing is not one specific image
Figure 4.5: Top row shows the mean shape of a cervical vertebra plus (light) and minus (dark) up to three standard deviations away in the direction of the first principal component. The four pictures represent the condition at the $s = 4$ steps that were required to fetch the query image: a) $s = 0$, b) $s = 1$, c) $s = 2$ and d) $s = 4$. Note the difference with figure 4.3 in that the four figures correspond with the four browsing steps, not the variations in the four principal component directions. Bottom row displays the best intermediate retrieval results: e) after the first query, f) after the second query g) after the third query and h) after the fourth query step. Note that the initial query is the result in d.

but a population, emphasizing exploitation rather than exploitation enables to define more carefully what is sought for. Images of the abnormal vertebrae ranked highest after each of the 4 browsing steps are illustrated in the bottom row of figure 4.5.

4.4 Automatic Image Classification

This far we used the expert classification of the vertebrae in the image repository to draw conclusions about its appearance and about the appearance of unknown but similar vertebrae. As a side issue we dedicate this section to classifying the images
automatically on the basis of their true segmentations.

We classify images in the repository on their commonly deviating characteristics from the population of normal vertebrae. For the $b$th image in the repository with feature function $f_b(v)$, the principal component score $g_b$ is computed with help of to the matrix of regression functions $B(v)$ defined in equation 4.9. This is done by solving

$$f_b(v) = g_bB(v)^T + \epsilon_b(v).$$

(4.33)

using least squares minimization of $\epsilon_b(v)$. This gives the optimal scores $g_1, ..., g_B$ corresponding to each feature function instance. As the score $g_b$ is computed according to the string model that is learned from the normal population, vertebrae in the image repository similar to the normal population will be grouped together, while vertebrae deviating from the normal population will form remote entities depending on the deviation.

To improve the discrimination ability of the scores $g_b$, we also take into account residual information. For each feature functions, its principal component score vector is augmented with the residual $\epsilon_b$ to obtain $g^*_b = [g_b, \epsilon_b]$. The residual is defined by

$$\epsilon_b = \frac{1}{N} \sum_{n=1}^{N} \left( \int_v \epsilon^v_b(v)dv - \frac{1}{B} \sum_{b=1}^{B} \int_v \epsilon^v_b(v)dv \right)$$

(4.34)

where $\epsilon^n_b(v)$ is the $n$th dimension of the residual function

$$\epsilon_b(v) = ||f_b(v) - g_bB(v)||^2.$$

(4.35)

The augmented scores $g^*_1, ..., g^*_B$ are clustered to obtain classes of scores with a similar deviations from the normal population. Unsupervised classification of the multidimensional score data is done by fitting a mixture of Gaussians with unconstrained covariance matrices and automatic choice of number of mixture components [55]. The result of clustering for $G = \{g_1, ..., g_B\}$ is given by

$$\mathcal{C}(G) = \{C_1, ..., C_C\}, C_i \subset G, \forall_{ij} : C_i \cap C_j = \emptyset.$$

(4.36)

Automatic image classification has two advantages. It releases the user from manually classifying images one by one. And, automatic clustering might reveal new classes of abnormal vertebrae only apparent when considering common deviations from the normal population in terms of multiple features.

## 4.5 Experiments and Results

The performance of the retrieval algorithm has been evaluated on the NHANES II cervical image data set [86]. As the data set includes a delineation of the vertebra boundaries in the form of point sets only, we obtain a continuous boundary representation by interpolating B-spline curves trough the points. B-spline curves are also used to represent feature functions, emanating from image and shape samples around the vertebra boundaries.
4.5.1 Experiments

For inductive learning we use a population of $M = 100$ normal cervical vertebra images. For commencing one browsing session we use one cervical vertebrae with lower osteophyte not present in the repository. This repository consists of $C = 5$ different images, a mixture of 45 normal vertebrae, 20 vertebrae with upper osteophyte, 68 with lower osteophyte and 49 with both upper and lower osteophyte, yielding a total of $B = 183$ images. The images for the different classes have been randomly selected out of the entire collection of 283 images. The composition of the images for learning and browsing is listed in table 4.1.

<table>
<thead>
<tr>
<th>Class</th>
<th>Total</th>
<th>Learning</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal cervical</td>
<td>145</td>
<td>100</td>
<td>45</td>
</tr>
<tr>
<td>Lower osteophyte</td>
<td>69</td>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>Upper osteophyte</td>
<td>20</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Lower and Upper</td>
<td>49</td>
<td>0</td>
<td>49</td>
</tr>
</tbody>
</table>

Table 4.1: Categories of cervical vertebrae and in what data set they have been used

We record vertebra boundaries by $N = 4$ features measured at 50 sample points along their B-spline representation. The first dimension $f^1(v)$ of the feature space $F$ is the distance between sampled x-coordinate values along the vertebra contour and the x-coordinate value of a single reference point on it. The second dimension $f^2(v)$ is the difference between sampled y-coordinate values and the y-coordinate value of the same reference point. The third dimension $f^3(v)$ is the local bending energy along the contour [79], measuring the curvature along the vertebra contour. The fourth dimension $f^4(v)$ is the gradient magnitude of the image along the vertebra contour, obtained by convolution of the image with the 2-dimensional first order fuzzy derivative at scale $\sigma = 4$ [44]. Hence, we have functions $f(v) = [f^1(v), f^2(v), f^3(v), f^4(v)]^T$ in 4-dimensional feature space.

**Automatic Image Classification**

In the first experiment we automatically classify images on the basis of their visual content. The aim is to investigate how accurately the images in the repository, described by $f_b(v)$, can be classified by considering their deviations from the normal cervical vertebra. We consider a case where $N = 2$ with only geometric features and a case where $N = 4$, with both shape and image features, i.e. $f_b(v) = [f_b^1(v), f_b^2(v), f_b^3(v), f_b^4(v)]^T$. The classification is done unsupervised using the clustering algorithm in [55], without specifying in advance the number and distribution of the true classifications. For each cluster we count the number of vertebrae assigned to the same class as the class assigned by experts [86].
Image Browsing and Retrieval

The goal of the second experiment is to determine whether the algorithm retrieves images that are relevant to the initial query. After definition of the unknown abnormal vertebra by segmentation of its image at the start of the browsing session, each subsequent iteration consists of selecting a new positive segmented image out of the top $\lambda = 10$ retrieved images. This selected image is used as the query in the next iteration. When the result is a screen full of pictures of negative images, the top ranked image is used as a negative example. Normally, the result of a query will offer a better alternative than the user did the last querying with and the algorithm will converge to the desired population.

To assess how relevant the retrieved images are to the initial query, we measure precision and recall [112], defined as

\[
\text{precision} = \frac{\text{No. relevant images retrieved}}{\text{Total No. images retrieved}}
\]

\[
\text{recall} = \frac{\text{No. relevant images retrieved}}{\text{Total No. relevant images in collection}}
\]

where high precision indicates that from all the images returned by a query, a large proportion of the images are relevant to the search (purity of retrieval). A high recall indicates that from all the images in the repository that are relevant to the query, a large number of these images are indeed returned (completeness of retrieval). In our case, the number of relevant images is the number of images that are of the same class as the initial query image, i.e. images by expert consensus classified as vertebra with lower anterior osteophyte.

We also address the question of how well the retrieval algorithm searches the space of solutions. We express the trade-off between exploration of the feature space and exploitation of previous results by

\[
\text{exploration} = \frac{\text{No. previously unretrieved relevant images}}{\text{Total No. of retrieved images}}
\]

\[
\text{exploitation} = \frac{\text{No. previously retrieved relevant images}}{\text{Total No. relevant images in collection}}
\]

where previously unretrieved images are images that where no part of the top $\lambda = 10$ best results in the history of the browsing session. In this context, exploration is the ability of the retrieval algorithm to search the feature space thoroughly; while exploitation refers to the algorithms ability to use the information about the feature space it has gained to narrow its future search.

4.5.2 Results

Automatic Classification Correctness

Results of image classification on the basis of all $N = 4$ shape and image features are given in table 4.2. A total of $C = 5$ separate clusters are found by the clustering
algorithm. All normal vertebrae are grouped into cluster 1. The vertebrae with lower osteophyte are predominantly found in cluster 3, whereas vertebrae with upper anterior osteophyte are also largely found in cluster 3. Cluster 2 primarily contains vertebrae with both lower and upper osteophyte. There is a clear distinction between classification correctness of normal and abnormal vertebrae. However, differences among abnormal vertebrae are less apparent. This is possibly due the fact that all vertebrae are explained with reference to the normal model, thereby possibly disregarding the characteristics of abnormal vertebrae. In comparison to classification on the basis of $N = 2$ shape features (data not shown here) application of both image and shape features improves the classification accuracy. Figure 4.6 shows individual images closest to centers of the clusters of table 4.2.

<table>
<thead>
<tr>
<th>Class</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal cervical</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lower osteophyte</td>
<td>17</td>
<td>3</td>
<td>31</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Upper osteophyte</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Lower and Upper</td>
<td>3</td>
<td>25</td>
<td>9</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 4.2: Automatic classification of the images in the repository on the basis of both shape and image features.

The use of residual information as defined in equation 4.35 contributes to a better classification than one without use of residuals (not shown here). However, as this information only signals when deviation is present not where, it is only suited to discriminate between, e.g. lower anterior osteophyte and lower-upper anterior osteophyte, for which the amount of deviation from the reference model is different. A better classification is expected on the basis of functional data, e.g. by functional discriminant analysis [95], rather than on the basis of principal component scores augmented with residual information.

The experiments with automatic image classification indicate there are no clear distinctions between features constituting the four classes of vertebrae. Rather the classes are characterized by common features, in addition to class-specific features, accentuating the importance of image browsing. For instance, if the user starts with an example image of lower osteophyte it is likely the algorithm retrieves normal vertebrae when that example happens to have many features in common with normal vertebrae. On the other hand, if the user can specifically indicate interest in lower parts of the vertebra but not in upper regions, by providing additional positive examples, the likelihood that only images from the lower osteophyte population will be returned increases. Forming a population of examples is expected to improve the discrimination ability of the algorithm and hence also that of retrieval accuracy.
Figure 4.6: Individual images from the repository closest to centers of the found clusters: a) normal cervical vertebrae corresponding to cluster 4, b) vertebra with lower anterior osteophyte corresponding to cluster 5, c) normal vertebra which happens to be closest to the center of cluster 3 and d) a vertebra with both lower and upper anterior osteophyte, closest to the center of cluster 2.

Precision and Recall for Retrieval

Figure 4.7 shows results for precision and recall. The results are the average of a total of 68 browsing sessions, each ending when the first $\lambda = 3$ of the ranked retrieved images are relevant. As the retrievals are ranked and the parameter $\lambda$ acts as stopping criteria, it is primarily the first part of the precision and recall graphs that provides valuable information. For the innovation rate $\gamma$ we have taken values between 0.5 and 1.0 as this range requires an acceptable number of iterations to convergence.

Figure 4.7a shows the precision graph, indicating the amount of relevant images shown on screen as a function of the number of retrievals. It can be seen that when a total of 50 images are retrieved about 45 percent of them are relevant for $\gamma = 0.5$ and $\gamma = 1.0$. The recall graph (not shown here) tells that at 50 retrievals about 30 percent of the relevant images in the repository are fetched. Concerning the innovation rate $\gamma$, the recall graph shows minor difference for the various values of $\gamma$, whereas the precision graph indicates that, especially for small number of retrievals, high values for $\gamma$ yield proportionally more precise results than lower values.

The precision versus recall graph shows for $\gamma = 0.5, \gamma = 0.7, \gamma = 0.9, \gamma = 1.0$ how precision decreases as increasingly large fractions of the collection are retrieved. It shows that to retrieve 50 percent of the relevant images from the repository, about 67 percent of the retrieved images will not be relevant, whereas for a recall of 10 percent, a precision of 75 percent is obtained. As concerns the influence of the innovation rate on the retrieval of relevant images, the graphs show that for $\gamma = 1.0$ in general the best performance is obtained.

For the vertebra application the precision and recall graphs tell that completeness of retrieval is difficult to control by population-based incremental learning due to
4.5. Experiments and Results

Figure 4.7: Average precision and recall values for 68 browsing sessions: a) recall graph b) precision versus recall graph.

diversity in the content of the images, while precision of retrieval is governed by the amount of focus on specific types of relevant images. From the above we conclude that image retrieval by exploration and exploitation on a population basis is promising when browsing an image repository with a specific target in mind.

Exploration and Exploitation for Browsing

Figure 4.8 shows the average exploration and exploitation graphs of 68 browsing sessions. Figure 4.8a illustrates the percentage of the retrieved relevant images that are not ranked in the top $\lambda = 10$ during the entire browsing session as a function of the number of retrieved images. It can be seen that up to approximately 10 retrievals all the retrieved images are suggested in the top $\lambda = 10$ at least once in the history of the browsing session. As expected most exploration is performed when $\gamma = 1.0$. For example, when a browsing session ends with 50 retrievals then for $\gamma = 1.0$, a percentage of 30 of the relevant retrieved images were not part of the top $\lambda = 10$ in the history of that session. For $\gamma = 0.5$ this is about 25 percent. We can state that high values for $\gamma$ truly assist the exploration of the image repository by investigating parts of the feature space not entered before.

The graph in figure 4.8b tells the percentage of the retrieved images that are suggested in the top $\lambda = 10$ as a function of the number of retrieved images. It can be seen that when a browsing session ends with 100 retrievals, then for $\gamma = 1.0$ a percentage of 5 of the relevant retrieved images have been part of the top $\lambda = 10$ in the history of the browsing session. For $\gamma = 0.5$ this is about 17 percent. Low values
for the innovation rate $\gamma$ hence allow to exploit previous results when browsing the image repository. We conclude that exploration and exploitation by population-based incremental learning allows to explicitly control the way of navigation through the image repository to reach a target population efficiently and accurately.

We also investigate the convergence of retrieval for various values of stopping criteria $\lambda$. The stopping criteria simply indicates the number of relevant images in the first $\lambda$ of the retrieved images. The graph in figure 4.9a shows the number of iterations required for convergence when $\gamma = 0.5, \gamma = 0.7, \gamma = 0.9$ and $\gamma = 1.0$. It can be seen that for $\lambda = 5$ a large number of iterations is required, especially for low values of the innovation rate $\gamma$. When the value of the innovation rate is large, on the other hand, even for $\lambda = 5$ a reasonable number of iterations is required for convergence. Hence, when aiming for retrieval of relevant images that highly resemble the query examples, a large number of browsing steps are required. The algorithm converges within an acceptable number of browsing steps when the emphasis is on retrieval of images that are less close to the example images, but are relevant to the query.

In determining the computational complexity of probabilistic image retrieval [123], we concentrate on scaling with size of the image repository. The graph in 4.9b gives the complexity when retrieving with $\gamma = 0.8$. The average number of iterations to convergence of ten browsing session is shown for increasing size of the image repository. As can be seen, for a small data set the average number of iterations is large. We partially attribute this to the fact that the query examples have features similar to a population not of interest for the query. In that case, the algorithm is not likely to

---

**Figure 4.8:** Average exploration and exploitation values for 68 browsing sessions: a) exploration graph and b) exploitation graph.
4.5. Experiments and Results

**Figure 4.9:** Average convergence and complexity measurements of 10 browsing sessions: a) convergence graph, b) complexity graph.

Fetch relevant images quickly. In fact, in the worst case it can oscillate. This problem diminishes with increasing size of the data set because then queries can be given far away from the overlapping areas. The complexity of the retrieval algorithm is difficult to determine due to the interplay of many of the components of the algorithm.

**Figure 4.10:** Illustration of a query image with retrieval results: a) initial query image is one of vertebra with lower anterior osteophyte, b) best matching image in repository after browsing, c) second best d) third best. Note that all the retrieved images are of vertebrae with lower anterior osteophyte.
Figure 4.10 shows example results of a browsing session. The initial query image is an image of a vertebra with lower anterior osteophyte. The best three matching images after browsing the retrieved images are shown. All three belong to the class of lower anterior osteophyte.

4.6 Discussion and Conclusion

We draw some important conclusions from the experiments we have conducted and the results thereof. In the first place, object-based image retrieval performs well if the segmentation model has the capacity to adapt to the user's concept. In comparing results of image retrieval by a fixed string segmentation model, with results obtained by an adaptive string segmentation model the latter reduces the minimal number of steps required for convergence. In the second place, image retrieval benefits greatly from exploring and exploiting the feature space by means of browsing retrieved images. The drawback of performing several browsing steps is paid back by an increase of the number of retrieved images that are relevant to the query. Thirdly, qualifying repository objects after each browsing step with respect to features that are recorded from the already browsed population with the same deviations from normal allows to quickly focus on the relevant images in the repository. In short, object-based image retrieval with help of an adaptive segmentation model as learned from a browsed population allows to learn the user's retrieval intention and recover relevant images in a limited number of steps.

A number of issues remain uninvestigated. It is expected that negative feedback improves the discrimination ability of the retrieval method in addition to dealing with local minima of the search space, e.g. when in response to user feedback the algorithm either returns exactly the same images as in a previous iteration or returns non-relevant images. The consequence of negative feedback to the exploration and exploitation of the feature space needs further study. Furthermore, in this application we have pre-selected a number of features for the definition of the object boundaries. A more precise definition of object-boundsaries requires studying a larger number of features. When the different object classes are known, as in our clinical example application, features can be studied separately for each class in order to build a dedicated string model for each abnormal object. This is expected to improve the discrimination ability of the proposed method and hence also that of retrieval accuracy.

We note that, although in this paper we applied the method for retrieval of vertebrae images, it can be used for any other object in any other application. In fact, the components of the proposed method can be considered as individual contributions that are not restricted to image retrieval. Functional data analysis [95] of a given set of multiple continuous boundary features gives detailed insight in the most important characteristics of objects and the natural variations therein. On top of this, such a statistical analysis is very much suited for multi-feature image segmentation and classification. It requires only a significant set of example images with delineations of objects. The population-based incremental learning technique originates from function optimization and has been used for a wide range application. This technique is
appealing because of its simplicity and because it generalizes easily. The string representation of objects is not a prerequisite for this algorithm since it relies on evolving a probabilistic model of a population rather than on the evolution of individuals.

In conclusion, the difference with other image retrieval approaches is that we do not aim at finding an object closest to a query object in a single step. Rather we try to incrementally build a distribution of example objects that explain the appearance of the target object to the degree that ranking images in the repository by their visual features yields relevant images. Construction of such a collection of examples is compelling because: a) the query object may be in an area with overlapping distributions, prohibiting any conclusions about the class it belongs to b) the segmentation model may not capture the query object properly, requiring more input objects to learn what is to be found in an image c) the feature extracted from segmentation may not be representative for the user’s retrieval intention, calling for a gradual optimization of features to consider. The combination of the string-based segmentation method with concepts borrowed from population-based incremental learning techniques [3] deals nicely with these problems. The proposed method works well for images of objects that are not easily defined by one feature, but require multiple features and multiple examples of these features to precisely capture the sought object.