

# On the Generation of Point Cloud Data Sets: Step One in the Knowledge Discovery Process

Andreas Holzinger<sup>1</sup>, Bernd Malle<sup>1</sup>, Marcus Bloice<sup>1</sup>, Marco Wiltgen<sup>1</sup>,  
Massimo Ferri<sup>2</sup>, Ignazio Stanganelli<sup>3</sup>,  
and Rainer Hofmann-Wellenhof<sup>4</sup>

<sup>1</sup> Research Unit Human-Computer Interaction, Institute for Medical Informatics,  
Statistics & Documentation,  
Medical University Graz, Austria

{a.holzinger,b.malle,m.bloice,m.wiltgen}@hci4all.at

<sup>2</sup> Department of Mathematics, University of Bologna, Italy

massimo.ferri@unibo.it

<sup>3</sup> Skin Cancer Unit IRCCS Istituto Tumori Romagna, Meldola, Italy

<sup>4</sup> Department of Dermatology, Graz University Hospital, Austria

**Abstract.** Computational geometry and topology are areas which have much potential for the analysis of arbitrarily high-dimensional data sets. In order to apply geometric or topological methods one must first generate a representative point cloud data set from the original data source, or at least a metric or distance function, which defines a distance between the elements of a given data set. Consequently, the first question is: How to get point cloud data sets? Or more precise: What is the optimal way of generating such data sets? The solution to these questions is not trivial. If a natural image is taken as an example, we are concerned more with the content, with the shape of the *relevant* data represented by this image than its mere matrix of pixels. Once a point cloud has been generated from a data source, it can be used as input for the application of graph theory and computational topology. In this paper we first describe the case for natural point clouds, i.e. where the data already are represented by points; we then provide some fundamentals of medical images, particularly dermoscopy, confocal laser scanning microscopy, and total-body photography; we describe the use of graph theoretic concepts for image analysis, give some medical background on skin cancer and concentrate on the challenges when dealing with lesion images. We discuss some relevant algorithms, including the watershed algorithm, region splitting (graph cuts), region merging (minimum spanning tree) and finally describe some open problems and future challenges.

**Keywords:** data preprocessing, point cloud data sets, dermoscopy, graphs, image analysis, skin cancer, watershed algorithm, region splitting, graph cuts, region merging, mathematical morphology.

## 1 Introduction and Motivation

Today we are challenged with complex, high-dimensional, heterogenous, and weakly-structured biomedical data sets and unstructured information from various sources [1]. Within such data, relevant structural or temporal patterns (“knowledge”) are often hidden, difficult to extract, and therefore not immediately accessible to the biomedical expert. Consequently, a major challenge is to interactively discover such patterns within large data sets. Computational geometry and algebraic topology may be of great help here [2], however, to apply these methods we need point cloud data sets, or at least distances between data entities. Point cloud data (PCD) sets can be seen as primitive manifold representation for use in algebraic topology [3]. For a rough guide to topology see [4].

A good example of a direct source for point clouds are 3D acquisition devices such as laser scanners, a recent low-cost commercial product being the Kinect device (see section 3). Medical images in nuclear medicine are also usually represented in 3D, where a point cloud is a set of points in the space, with each node of the point cloud characterized by its position and intensity (see section 3 and 5). In dimensions higher than three, point clouds (feature vectors) can be found in the representation of high-dimensional manifolds, where it is usual to work directly with this type of data [5].

Some data sets are naturally available as point clouds, for example protein structures or protein interaction networks, where techniques from graph theory can be directly applied [6].

Despite the fact that naturally occurring point clouds do exist, a concerted effort must focus on how to get representative point cloud data sets from raw data. Before continuing, and for clarification purposes, some key terms are defined in the next section. This is followed by discussing natural point clouds in section 3 as well as the case of text documents in section 4, before examining the case of medical images, and in particular dermatological images, in section 5. We first introduce some dermatological image sources, describe shortly some problems facing the processing of such images, and present some related work, as well as relevant algorithms. Finally, we discuss open problems and provide an outline to future research routes in sections 6 and 7, respectively.

## 2 Glossary and Key Terms

*Point clouds:* are finite sets equipped with a family of *proximity* (or *similarity measure*) functions  $sim_q: S^{q+1} \rightarrow [0, 1]$ , which measure how “close” or “similar”  $(q + 1)$ -tuples of elements of  $S$  are (a value of 0 means totally different objects, while 1 corresponds to essentially equivalent items).

*Space:* a set of points  $a_i \in S$  which satisfy some geometric postulate.

*Topology:* the study of shapes and spaces, especially the study of properties of geometric figures that are not changed by continuous deformations such as stretching (but might be by cutting or merging) [7], [8].

*Topological Space:* A pair  $(X, T)$  with  $X$  denoting a non-empty set and  $T$  a collection of subsets of  $X$  such that  $\emptyset \in T$ ,  $X \in T$  and arbitrary unions and finite intersections of elements of  $T$  are also  $\in T$ .

*Algebraic Topology:* the mathematical field which studies topological spaces by means of algebraic invariants [9].

*Topological Manifold:* A topological space which is locally homeomorphic (has a continuous function with an inverse function) to a real  $n$ -dimensional space (e.g. Euclidean space) [10].

*Distance:* Given a non-empty set  $S$ , a function  $d : S \times S \rightarrow \mathbb{R}$  such that for all  $x, y, z \in S$  (i)  $d(x, y) \geq 0$ , (ii)  $d(x, y) = 0 \iff x = y$ , (iii)  $d(x, y) = d(y, x)$ , and (iv)  $d(x, z) \leq d(x, y) + d(y, z)$ .

*Metric space:* A pair  $(S, d)$  of a set and a distance on it. Every metric space is automatically also a topological space.

*Computational geometry:* A field concerned with algorithms that can be defined in terms of geometry (line segments, polyhedra, etc.) [11].

*Supervised Learning:* Method within Machine Learning that uses labeled training data to develop an accurate prediction algorithm. Let  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  be  $n$  training samples with  $x_1 \dots x_n$  being the predictor variables and  $y_1 \dots y_n$  the labels, we want a function  $g : X \rightarrow Y$  such that a cost function (usually the difference between predicted values  $g(x)$  and  $y$ ) is minimized.

*Unsupervised Learning:* Method in machine learning which is used to group similar objects together, e.g. points within geometric groups or objects of similar properties (color, frequency). No labeled training data is used.

*Optimization:* is the selection of cluster a best element (with regard to some criteria) from some set of available alternatives.

*Classification:* Identification to which set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known.

*Clustering:* Grouping a set of objects in such a way that objects in the same group (cluster) are more similar to each other than to those in other groups (clusters).

*Feature:* A measurable property of an object (e.g. the age of a person).

*Feature Vector:* A collection of numerical features interpreted as the dimensional components of a (Euclidean) vector.

*Vector space model:* Approach whose goal is to make objects comparable by establishing a similarity measure between pairs of feature vectors (Euclidean distance, cosine similarity etc.). The space spanned by all possible feature vectors is called the feature space.

*Voronoi region:* Given a set of points in a metric space  $p_1, \dots, p_n$ , a Voronoi diagram erects regions around a point  $p_i$  such that all points  $q$  within its region are closer to  $p_i$  than to any other point  $p_j$  [12].

*Delaunay triangulation:* Given a set of points in a plane  $P = p_1, \dots, p_n$ , a Delaunay triangulation separates the set into triangles with  $p_i \in P$  as their corners, such that no circumcircle of any triangle contains any other point in its interior.

*Minimum Spanning Tree:* Given a graph  $G = (V, E, \omega)$  with  $V$  being the set of vertices,  $E$  being the set of edges and  $\omega$  being the sets of edge weights, a Minimum Spanning tree is the connected acyclic subgraph defined by the subset  $E' \subseteq E$  reaching all vertices  $v \in V$  with the minimal sum of edge weights possible.

### 3 The Case for Natural Point Clouds

A prototypical example of natural point clouds are the data produced by 3D acquisition devices (Figure 1, Left), such as laser scanners [13]. Methods for the extraction of surfaces from such devices can roughly be divided into two categories: those that segment a point cloud based on criteria such as proximity of points and/or similarity of locally estimated surface normals, and those that directly estimate surface parameters by clustering and locating maxima within a parameter space; the latter is more robust, but can only be used for simple shapes such as planes and cylinders that can be described by only a few parameters [14]. A recent low-cost example is the Kinect (Figure 1, Center) device [15]. This sensor is particularly interesting as such devices will continue to gain popularity as their prices drop while at the same time becoming smaller and more powerful and the open source community will promote its use [16]. Such sensors have the potential to be used for diverse mapping applications; however, the random error of depth measurement increases with increasing distance to the sensor, and ranges from a few millimeters up to about four centimeters at the maximum range of the Kinect device [17]. Some recent examples demonstrate the potential of this sensor for various applications, where high precision is not an issue, e.g. in rehabilitation exercises monitoring [18] or in health games [19].

It seems reasonable to assume the presence of 3D-scanners within mobile devices in the not-so-distant future, which in combination with faster, more

powerful algorithms and advances in software engineering could potentially transform each smartphone into a mobile medical laboratory and mobile wellness center [20]. Although applications in this area will likely not be an adequate substitution for the work of trained professionals, it would help reduce data pre-processing time and could make some hospital visits for purely diagnostic purposes a thing of the past, consequently help to tame the worldwide exploding health costs.

Medical images, e.g. in nuclear medicine, are usually also represented in 3D, following the same principle (Figure 1), where a point cloud is a set of points in  $\mathbb{R}^3$ , whose vertices are characterized by their position and intensity. The density of the point cloud determines the resolution, and the reconstructed volume, which in general could be of any resolution, size, shape, and topology, is represented by a set of non-overlapping tetrahedra defined by the points. The intensity at any point within the volume is defined by linearly interpolating inside a tetrahedron from the values at the four nodes that define such a tetrahedron, see [21] for more details and see [22] for some basic principles.

Some data sets are "naturally" available as point clouds, which is convenient as  $n$ -dimensional point clouds can easily be mapped into graph data structures by defining some similarity functions to pairs of nodes (e.g. the Euclidean distance, however a multitude of methods are available) and assigning that similarity to edges between them. Examples of this include protein structures or protein interaction networks (Figure 1, Right), where techniques from graph theory can be applied [6].



**Fig. 1.** Left: A 3D scan of Bernd Malle taken in 1998 by a stationary device worth around EUR 100,000. Center: 3D scan taken in 2013 by a Microsoft Kinect device worth EUR 200 (Source: <http://www.kscan3d.com/>). Right: Protein-protein interaction network (Source: <http://www.pnas.org/>).

## 4 The Case of Text Documents

Based on the vector space model, which is a standard tool in text mining [23], a collection of text documents (aka corpus) can be mapped into a set of points (vectors) in  $\mathbb{R}^n$ . Each word can also be mapped into vectors, resulting in a very

high dimensional vector space. These vectors are the so-called term vectors, with each vector representing a single word. If there, for example, are  $n$  keywords extracted from all the documents then each document is mapped to a point (*term vector*) in  $\mathbb{R}^n$  with coordinates corresponding to the weights. In this way the whole corpus can be transformed into a point cloud set. Usually, instead of the Euclidean metric, using a specialized similarity (proximity) measure is more convenient. The *cosine similarity measure* is one example which is now a standard tool in text mining, see for example [24]. Namely, the cosine of the angle between two vectors (points in the cloud) reflects how “similar” the underlying weighted combinations of keywords are. By following this approach, methods from computational topology may be applied [25], which offers a lot of interesting research perspectives.

## 5 The Case of Medical Images

### 5.1 Some Fundamentals of Digital Images

**Dermoscopy.** The dermoscopy, aka epiluminescence microscopy (ELM), is a non-invasive diagnostic technique and tool used by dermatologists for the analysis of pigmented skin lesions (PSLs) and hair, that links clinical dermatology and dermatopathology by enabling the visualization of morphological features otherwise not visible to the naked eye [26]. Digital dermoscopy images can be stored and later compared to images obtained during the patient’s next visit for melanoma and non-melanoma skin cancer diagnosis. The use of digital dermoscopes permitted the documentation of any examinations in the medical record [27] [28].

Skin and pathology appearance varies with light source, polarization, oil, pressure, and sometimes temperature of the room, so it is important that the examination and documentation be performed in a standardized manner. To do so, some of the most modern spectrometers use an adjustable light source which adjusts according to the room light to try to mimic the “daylight” spectrum from a standardized light source.

Although images produced by polarised light dermoscopes are slightly different from those produced by a traditional skin contact glass dermoscope, they have certain advantages, such as vascular patterns not being potentially missed through compression of the skin by a glass contact plate. Dermoscopy only evaluates the down level of papillary dermis, leaving pathologies in the reticular dermis unseen. Amelanotic melanoma is missed with this method and high pigmented lesions can also hide structures relevant for the diagnosis. A negative surface exam is no guarantee that there is no pathology. In case of doubt a biopsy and experienced clinical judgment is required [29].

**Confocal Laser Scanning Microscopy.** Reflectance confocal microscopy (RCM) allows non-invasive imaging of the epidermis and superficial dermis. Like dermoscopy, RCM acquires images in the horizontal plane (en face), allowing

assessment of the tissue pathology and underlying dermoscopic structures of interest at a cellular-level resolution [30].

The confocal image uses a low-power laser and special optics to magnify living cells at approximately 1,000 times zoom. Confocal images are created by the natural refractive difference within and between cells. Melanin is highly refractive and results in brighter images. The confocal microscope captures the images in three dimensions within the top layers of the skin. Imaging each lesion takes between 5 and 10 minutes. A tissue ring is attached with medical adhesive to hold the skin stable and the laser tracks through the lesion in three dimensions to create vertical and horizontal maps of the cell fields. There is no pain or scarring in this non-invasive procedure [29].

The application of a wide array of new synthetic and naturally occurring fluorochromes in confocal microscopy has made it possible to identify cells and sub-microscopic cellular components with a high degree of specificity amid non-fluorescing material. In fact the confocal microscope is often capable of revealing the presence of a single molecule. Confocal microscopy offers several advantages over conventional widefield optical microscopy, including the ability to control depth of field, elimination or reduction of background information away from the focal plane (which leads to image degradation), and the capability to collect serial optical sections from thick specimens, making possible multi-dimensional views of living cells and tissues that include image information in the  $x$ ,  $y$ , and  $z$  dimensions as a function of time and presented in multiple colours (using two or more fluorophores). The temporal data can be collected either from time-lapse experiments conducted over extended periods or through real time image acquisition in smaller frames for shorter periods of time. A concise overview on biological image analysis can be found here [31].

**Total-Body Photography.** Total body photography (TBP) is a diagnostic technique where a series of high resolution digital photographs are taken from head to toe of the patients skin for active skin cancer surveillance [32]. A photographic baseline of the body is important when attempting to detect new lesions or changes in existing lesions in patients with many nevi and create a pigment lesion mapping of the entire body. Changes in moles can be in the form of size, shape and colour change and it can also be useful for other conditions as psoriasis or eczema.

The main advantages of total body photography are that it reduces unnecessary biopsies, and melanomas are often caught at a much earlier stage. A recent approach is Gigapixel Photography (GP), which was used to capture high-res panoramas of landscapes; recent developments in GP hardware have led to the production of consumer devices (see e.g. [www.GigaPan.com](http://www.GigaPan.com)). GP has a one billion pixel resolution capacity, which is 1000 times higher than TBP, and therefore has a lot of potential for dermatology use [33].

## 5.2 Point Cloud Data Sets

To sum up at this point, let us define in accordance with [34]:

- **Multivariate dataset** is a data set that has *many dependent variables* and they might be correlated to each other to varying degrees. Usually this type of dataset is associated with *discrete data models*.
- **Multidimensional dataset** is a data set that has *many independent variables* clearly identified, and one or more dependent variables associated to them. Usually this type of dataset is associated with *continuous data models*.

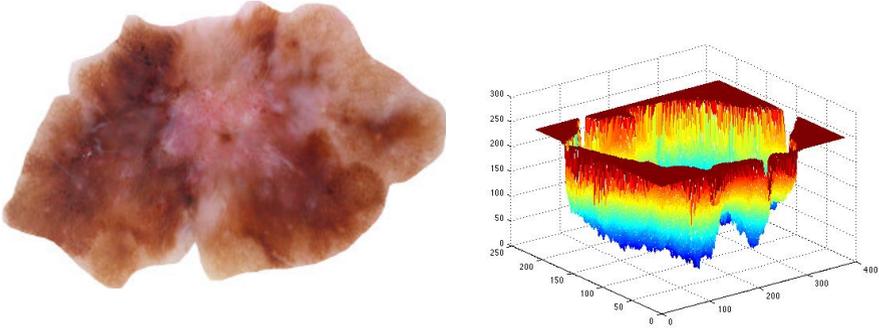
In other words, every data item (or object) in a computer is represented (and therefore stored) as a set of features. Instead of the term features we may use the term dimensions, because an object with  $n$  features can also be represented as a multidimensional point in an  $n$ -dimensional space. Dimensionality reduction is the process of mapping an  $n$ -dimensional point, into a lower  $k$ -dimensional space, which basically is the main challenge in visualization .

The number of dimensions can sometimes be small, e.g. simple 1D data such as temperature measured at different times, to 3D applications such as medical imaging, where data is captured within a volume. Standard techniques like contouring in 2D, and isosurfacing and volume rendering in 3D, have emerged over the years to handle these types of data. There is no dimension reduction issue in these applications, since the data and display dimensions essentially match.

One fundamental problem in analysing images via graph theoretical methods is when first translating them into a point cloud. While pixels in images naturally have some coordinates in 2D, their colour value as well as relation to pixels around them is not encoded within those coordinates. Thus, some transformation of the 2D image into a higher-dimensional space has to occur as a first step. This, however, entails many problems such as inadvertently modelling artefacts or ‘inventing’ information that is not contained in the image. The following gives an example of a simple 2D to 3D transform of a melanoma image (Figure 2).

## 5.3 Two Examples of Creating Point Clouds

The functional behaviour of a genome can be studied by determining which genes are induced and which genes are repressed in a cell during a defined snapshot. The behaviour can change in different development phases of the cell (from a stem cell to a specialized cell), in response to a changing environment (triggering of the gene expression by factor proteins with hormonal function) or in response to a drug treatment. The microarray technology makes it possible to explore gene expression patterns of entire genomes (a recent work from cancer research can be found in [35]). Technically, a microarray is usually a small glass slide (approximately 2.0 cm  $\times$  2.0 cm) covered with a great number (20,000 or more) of precisely placed spots. Each spot contains a different single stranded DNA sequence fragment: the gene probe. A microarray experiment is done as follows: From reference and test tissue samples, mRNA is isolated and converted into cDNA. The cDNAs are labelled green (reference) and red (test). The cDNA



**Fig. 2.** Simple 2D-3D transform using the mesh function built into MATLAB

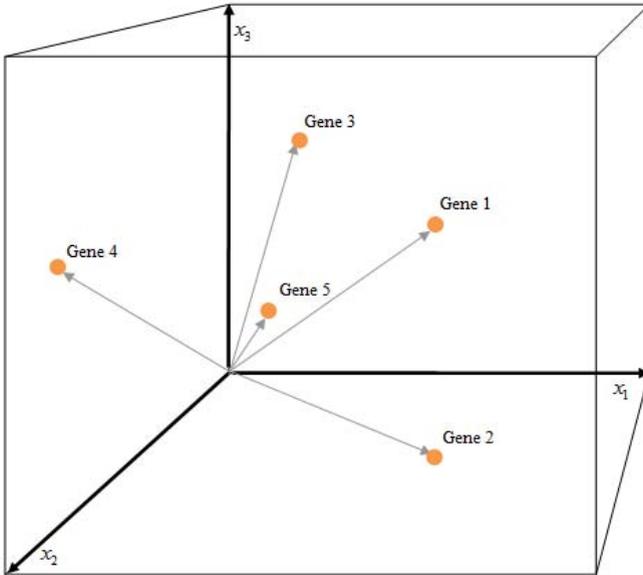
samples are mixed together and incubated with the probe on the microarray. The location and intensities of the fluorescent dyes are recorded with a scanner. Using the ratios of the dyes, a red spot indicates induced gene activity within the test probe; a green spot shows repressed gene activity in the test probe and a yellow spot indicates that there is no change in the gene activity level in the two probes. The amount of data resulting from microarray experiments is very big and too complex to be interpreted manually by a human observer. Machine learning algorithms extract from a vast amount of data the information that is needed to make the data interpretable. The gene expression pattern of the gene  $y_n$  along  $P$  experiments is described by a vector:

$$y_n = (x_{n1}, x_{n2}, \dots, x_{nk}, \dots, x_{nP})$$

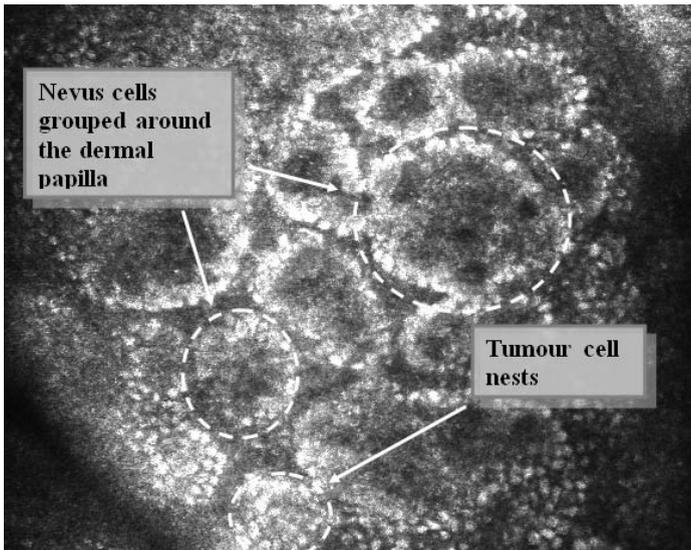
where  $x_{nk}$  is the expression value of the gene during the experiment number  $k$ . The genes can be geometrically interpreted as a point cloud in a  $P$ -dimensional space (Figure 3).

In the diagnosis of CLSM views of skin lesions, architectural structures at different scales play a crucial role. The images of benign common nevi show pronounced architectural structures, such as arrangements of nevi cells around basal structures and tumour cell nests (Figure 4).

The images of malign melanoma show melanoma cells and connective tissue with few or no architectural structures. Features based on the wavelet transform have been shown to be particularly suitable for the automatic analysis of CLSM images because they enable an exploration of images at different scales. The multi resolution analysis takes scale information into consideration and successively decomposes the original image into approximations (smooth parts) and details. That means, through the wavelet transformation, the two-dimensional image array is split up into several frequency bands (containing various numbers of wavelet coefficients), which represent information at different scales. At each scale the original image is approximated with more or fewer details. The frequency bands, representing information at a large scale, are labelled with



**Fig. 3.** Every gene is represented by a point in a  $P$ -dimensional space, which is built by the  $P$  experiments (for example:  $P$  different kinds of tissue). The position of the point is determined by the expression values on each axis of the coordinate system.



**Fig. 4.** Nevi cell arrangement and tumour cell nests

low indices and the frequency bands representing successively decreasing scales are labelled with higher indices. Then the architectural structure information in the CLSM images is accumulated along the energy bands (from course to fine). Therefore the wavelet transformation allows the analysis of a given texture by its frequency components. In wavelet texture analysis, the features are mostly derived from statistical properties of the resulting wavelet coefficients inside the frequency bands. Then the tissue textures, in the CLSM images, are represented by feature vectors, as for example:

$$\mathbf{x}_n = \left( F_{\text{STD}}^n(i) \right); i = 0, \dots, N$$

Whereby  $N$  is the number of frequency bands. The index  $n$  refers to the  $n$ -th image.  $F_{\text{STD}}^n(i)$  represents a statistical property of the wavelet coefficients in the  $i$ -th frequency band. From an ensemble of images results a point cloud of different feature vectors in the feature space.

**Data Set Example:** A relatively recent development is the creation of the UCI Knowledge Discovery in Databases Archive available at <http://kdd.ics.uci.edu>. This contains a range of large and complex datasets as a challenge to the data mining research community to scale up its algorithms as the size of stored datasets, especially commercial ones, inexorably rises [36].

## 5.4 Graphs in Image Analysis

The idea of using graph theoretic concepts for image processing and analysis goes back to the early 1970's. Since then, many powerful image processing methods have been formulated on pixel adjacency graphs. These are graphs whose vertex set is the set of image elements (pixels), and whose edge set is determined by an adjacency relation among the image elements.

More recently, image analysis techniques focus on using graph-based methods for segmentation, filtering, clustering and classification. Also, graphs are used to represent the topological relations of image parts.

**Definition 1 (Graph).** A graph  $G = (V, E)$  is given by a finite set  $V$  of elements called vertices, a finite set  $E$  of elements called edges, and a relation of incidence, which associates with each edge  $e$  an unordered pair  $(v_1, v_2) \in V \times V$ . The vertices  $v_1$  and  $v_2$  are called the end vertices of  $e$ .

**Definition 2 (Planar Graph, Embedded Graph).** A graph is said to be **planar** if it can be drawn in a plane so that its edges intersect only at its end vertices. A graph already drawn in a surface  $S$  is referred to as **embedded in  $S$**  [37].

## 5.5 Medical Background

**Skin Cancer** is still the most common and most increasing form of human cancer worldwide. Skin cancer can be classified into melanoma and non-melanoma and although melanomas are much less common than non-melanomas, they account for the majority of skin cancer mortality. Detection of malignant melanoma in its early stages considerably reduces morbidity and mortality and may save hundreds of millions of Euros that otherwise would be spent on the treatment of advanced diseases. If cutaneous malign melanoma can be detected in its early stages and removed, there is a very high likelihood that the patient will survive.

However, melanomas are very complex and a result of accumulated alterations in genetic and molecular pathways among melanocytic cells, generating distinct subsets of melanomas with different biological and clinical behavior. Melanocytes can proliferate to form nevi (common moles), initially in the basal epidermis [38]. A melanoma can also occasionally simply look like a naevus.

Image analysis techniques for measuring these features have indeed been developed. The measurement of image features for the diagnosis of melanoma requires that lesions first be detected and localized in an image. It is essential that lesion boundaries are determined accurately so that measurements, such as maximum diameter, asymmetry, irregularity of the boundary, and color characteristics, can be accurately computed. For delineating lesion boundaries, various image segmentation methods have been developed. These methods use color and texture information in an image to find the lesion boundaries [39].

## 5.6 Challenges

Basic difficulties when dealing with such lesions include:

1. **Morphology is not enough**

Melanomas can sometimes appear like naevi. This suggests relying on follow-ups and to perhaps prefer sensitivity to specificity.

2. **Detail Level**

Medical doctors are understandably fond of details, whereas preprocessing often needs to blur images together with noise.

3. **Diversity**

Especially in dermoscopy there is a great variety of established criteria to describe melanocytic and non melanocytic lesions [40].

4. **Segmentation**

Segmentation is one of the main hurdles in lesion analysis, as a good segmentation of different skin lesions is crucial for total body imaging. It is also seen as a problem by dermatologists themselves [41]: There has been research on interoperator and *intraoperator* differences in segmentation by hand of one and the same lesion.

5. **Noise**

Having said that, it is a requirement to split the lesion from the background. This is even more problematic with people of darker complexion. A further



**Fig. 5.** A naevus (left) and a melanoma (right)[42]

problem is hair: the pragmatic solution is to physically remove any hair using a razor. But it is much better to eliminate it (or to “ignore” it) at the image level, for example through the use of the Hough transform [43].

The **Hough transform** is a method for detecting curves by exploiting the duality between points on a curve and the parameters of that curve, hence it is well-suited for the detection of arbitrary shapes, and therefore ideal for removing hair. The method is robust to partial deformation in shape and very tolerant to noise and can detect multiple occurrences of a shape in the same region, however, it requires a lot of memory and computational power [44].

The most optimal segmentations we obtained were through the Mumford-Shah functional, but it requires much processing time [45]. It is therefore better to rely on a cleverly engineered mixture of morphological operations and thresholding. A big issue here is in interactivity, because exactly here the expert end user could come into play, by making her/him either simply to accept or reject a segmentation or even initialize it or modify it (see below Interaction with the user).

## 6. Diagnostic Criteria

Dermatologists trust the following criteria:

- A: Asymmetry
- B: Boundary (border irregularity)
- C: Colour (variegation and uneven distribution)
- D: Diameter (greater than 6 mm)
- E: Elevation (Alternatively: Evolution)

Moreover in patients with many nevi or other skin lesions this simplified algorithm is not sufficient to diagnose such lesions correctly. Experience, comparison of multiple lesions, and follow-up information is crucial to come to a correct diagnosis. At this point one may ask how to make this procedure at least partially automatic, and persistent homology is certainly one approach, as we shall see.

## 7. Interaction with the User

Interaction in time is an important issue. Although it is unreasonable to expect “real time” outputs, a procedure in the order of minutes is a far too long time for a medical doctor and also for a patient. A processing time of approximately 2 minutes, which is usual considering the aforementioned criteria, requires that something be put on the screen, showing that the computer has not frozen and that something is actually happening.

Moreover the output must be understandable. Therefore, a trade-off between richness and simplicity of information is required. One possibility is to have two (hidden) classifiers, one “pessimistic” algorithm (tuned to high sensitivity) and one “optimistic” algorithm (high specificity). This, however, can result in three possible outputs: High risk (both classifiers agreeing on melanoma), medium risk (disagreeing), and low risk (both agreeing on naevus). This approach is certainly not satisfactory for the present purposes.

## 8. Representation

On the strictly technical side, one can simply represent the images as graphs with pixels as vertices, and 4-neighbours as adjacent vertices. Of course, much more elaborate methods have been developed, which shall be discussed further in the following sections.

## 5.7 Related Work

De Mauro, Diligenti, Gori & Maggini [46] in 2003 presented a very relevant piece of work: they proposed an approach based on neural networks by which the retrieval criterion is derived on the basis of learning from examples. De Mauro et al. used a graph-based image representation that denoted the relationships among regions in the image and on recursive neural networks which can process directed ordered acyclic graphs. This graph-based representation combines structural and sub-symbolic features of the image, while recursive neural networks can discover the optimal representation for searching the image database. Their work was presented for the first time at the GBR 2001 conference in Ischia and the authors subsequently expanded it for a journal contribution.

Bianchini (2003) [47] reported on the computationally difficult task of recognizing a particular face in a complex image or in a video sequence, which humans can simply accomplish using contextual information. The face recognition problem is usually solved having assumed that the face was previously localized, often via heuristics based on prototypes of the whole face or significant details. In their paper, they propose a novel approach to the solution of the face localization problem using recursive neural networks. In particular, the proposed approach assumes a **graph-based representation of images** that combines structural and subsymbolic visual features. Such graphs are then processed by recursive neural networks, in order to establish the eventual presence and the position of the faces inside the image.

Chen & Freedman (2011) [48] reported on an alternative method in the pre-processing stage: In cortex surface segmentation, the extracted surface is required to have a particular topology, namely, a two-sphere. The authors presented a

novel method for removing topology noise of a curve or surface within the level set framework, and thus produce a cortical surface with the correct topology. They defined a new energy term which quantifies topology noise and showed how to minimize this term by computing its functional derivative with respect to the level set function. This method differs from existing methods in that it is inherently continuous and not digital; and in the way that our energy directly relates to the topology of the underlying curve or surface, versus existing knot-based measures which are related in a more indirect fashion.

## 5.8 Relevant Algorithms

**The Watershed Algorithm** is a popular tool for segmenting objects whose contours appear as crest lines on a gradient image as it is the case with melanomas. It associates to a topographic surface a partition into catchment basins, defined as attraction zones of a drop of water falling on the relief and following a line of steepest descent [49].

Each regional minimum corresponds to such a catchment basin. Points from where several distinct minima may be reached are problematic as it is not clear to which catchment basin they should be assigned. Such points belong to watershed zones, which may be thick. Watershed zones are empty if for each point, there exists a unique steepest path towards a unique minimum. Unfortunately, the classical watershed algorithm accepts too many steep trajectories, as they use neighborhoods which are too small for estimating their steepness. In order to produce a unique partition despite this, they must make arbitrary choices that are out of the control of the user. Finally, their shortsightedness results in imprecise localizations of the contours.

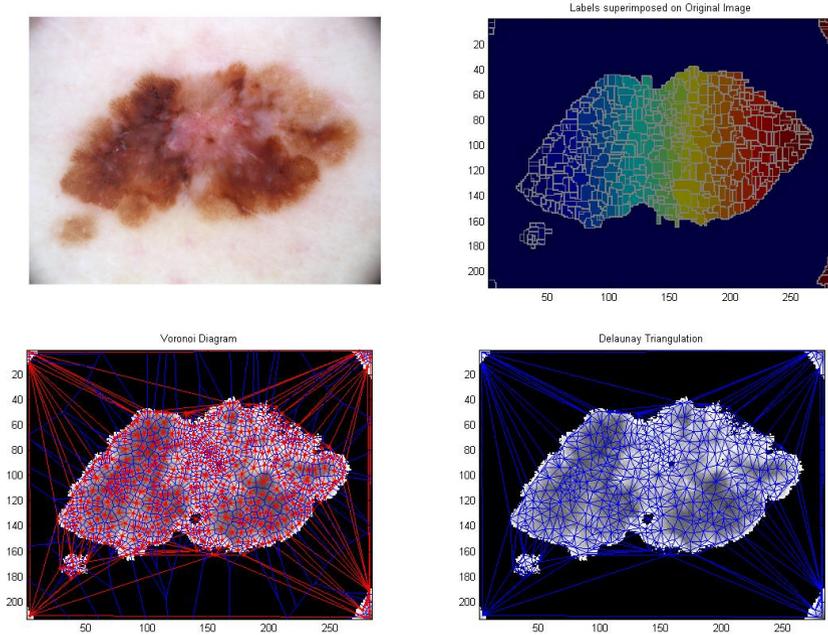
We propose an algorithm without myopia, which considers the total length of a trajectory for estimating its steepness; more precisely, a lexicographic order relation of infinite depth is defined for comparing non ascending paths and choosing the steepest. For the sake of generality, we consider topographic surfaces defined on node weighted graphs. This allows us to easily adapt the algorithms to images defined on any type of grid in any number of dimensions. The graphs are pruned in order to eliminate all downwards trajectories which are not the steepest. An iterative algorithm with simple neighborhood operations performs the pruning and constructs the catchment basins. The algorithm is then adapted to gray tone images. The neighborhood relations of each pixel are determined by the grid structure and are fixed; the directions of the lowest neighbors of each pixel are encoded as a binary number. In that way, the graph may be recorded as an image. A pair of adaptive erosions and dilations prune the graph and extend the catchment basins. As a result, one obtains a precise detection of the catchment basin and a graph of the steepest trajectories [50].

Note: Stable image features, such as SIFT or MSER features, can also be taken to be the nodes of the graph.

The watershed segmentation is a **regionbased technique** making use of image morphology; a classic description can be found in [51]. It requires the

selection of markers (“seed” points) interior to each object of the image, including the background as a separate object.

The markers are chosen by a human expert who takes into account the application-specific knowledge of the objects. Once the objects are marked, they can be grown using a morphological watershed transformation (Figure 6) [52].



**Fig. 6.** Result of applying a watershed transform to an image of relatively distinguished regions. The resulting segmentation (and thus vertices for the output graph) corresponds well to the overall shape of the image and represents regions of about equal size.

## 5.9 Region Splitting (Graph Cuts)

Understanding the original image as a graph consisting of one large, connected component, the goal of region splitting is to obtain a graph  $G(V, E)$  with a number of vertices ( $|V|$ ) significantly smaller than the number of input pixels ( $|V| \ll n$ ). In order to achieve this we have to group certain areas consisting of varying amounts of pixels together. This can be done via a partition of the image, with a partition being defined as a subgraph ( $G'(V, E')$ ) of the original graph with the set of vertices being the same as in the original and the set of edges being a strict subset of the original set ( $E' \subset E$ ) (one must remove edges in order to separate formerly connected components). This separation occurs recursively until a cutting threshold is obtained for all remaining connected components,

which are then interpreted as regions or superpixels (-voxels) represented by some additionally extracted information, stored as a feature vector for the individual partition.

### 5.10 Region Merging (Minimum Spanning Tree)

This is essentially the opposite from the method just mentioned, in that the input image is considered as a set of pixels, each constituting its own region. The goal is to merge regions based on a (dis-)similarity measure. Felzenswalb (2004) [53] proposed an algorithm which in effect defines one numerical figure representing the internal similarity of a region, and a second figure representing the dissimilarity between two adjacent regions. In short, the approach works like this:

$$\text{Int}(C) = \max_{e \in \text{MST}(C,E)} \omega(e)$$

is the internal region similarity figure, given by the maximum edge weight of the regions MST (Minimum Spanning Tree).

$$\text{Dif}(C1, C2) = \min_{v_i \in C1, v_j \in C2, (v_i, v_j) \in E} \omega(v_i, v_j)$$

denotes any two regions' dissimilarity figure, given by the minimum edge weight connecting them.

Finally,

$$D(C1, C2) = \begin{cases} \text{true} & \text{if } \text{Dif}(C1, C2) > \text{MInt}(C1, C2) \\ \text{false} & \text{otherwise} \end{cases}$$

determines if two regions should be merged, based on the relation of their inter-region dissimilarity and minimum respective internal similarities.

As per the region splitting approach, once no further regions can be merged, the final image partition is obtained.

## 6 Open Problems

### 6.1 Medical Problems

One of the greatest problems in skin cancer screening is to select the right lesion for further investigation. An adult person has anywhere between 20 and one hundred different lesions. The segmentation and recognition of suspicious lesions, which need further investigation by dermoscopy or RCM or another procedure, is of utmost importance.

Furthermore the differentiation of physiologic changes from malignant changes in a lesion is a great challenge for the dermatologist. The same is true for the validation of benign and malignant criteria in one lesion. The question is, does a small part of the lesion showing criteria of malignancy justify an excision or not?

## 6.2 Graphs from Images

In implementing and testing different techniques for graph extraction out of medical image data several areas of consideration have arisen.

**Is a Graph a Good Representation of an Image?** This is logical as every image consists of pixels that share some form of collocation with one another, may it be geometrical neighborhoods or distances in some feature space. Secondly, image representation through graphs is already used by several segmentation algorithms, as the above sections have sufficiently discussed. The difference between our approach and the aforementioned methods is that the former are treating the raw structure of the image as the input graph to their algorithm, whose output then is a general segmentation. This work however intends to produce a graph representation of the image as its output for further use, while it may or may not use a graph based algorithm to compute it.

**Why Compute another Graph?** One could argue that every image in pixel form (there are other representations like wavelets used in JPG) already contains an implicit graph. While this is certainly true, an image of several megapixels would translate to a graph containing millions of vertices ( $n$ ) and (given a  $k$ -neighborhood for each pixel)  $m = k * n$  edges. This input size is clearly too large for any algorithm of polynomial runtime complexity, especially if it is intended to be used on standard desktop computers or even mobile devices. It is thus imperative to reduce the number of vertices by first applying some form of segmentation or clustering.

**Can a Reliable Graph Be Extracted from One Image Alone?** Another interesting question is how well a (2D) image represents a surface topography at all. Usually the only pieces of information contained in an image are the coordinates of its pixels plus their corresponding color values. The latter (after a transform to an intensity value) is typically interpreted as the height of its pixel, thereby transforming the image to a topographic map. This information however might be imprecise due to light conditions at the time of photography, hardware inaccuracies, angle of the recording device etc., leading to artifacts and thus misrepresentation. The only solution to this problem would be to take several images in a sequence over time, from different angles, or applying a different image taking technology (3D or radar scanning) altogether.

Based on this, a topological analysis of a graph extracted and merged from several images (sources) might reveal information not contained in a single image, while avoiding the incorporation of the same artifacts or inaccuracies that a single input source might contain.

**Is Image Segmentation the Pertinent Approach in Our Case.** In traditional applications the goal of segmenting an image is mostly object recognition

or detection, either as an unsupervised grouping of areas belonging together or by matching a template representation to an area within the image, often considering different scales, angles or other deviations. The output segmentation of this class of tasks consists of a small set of image regions, representing either the locations of potential object matches or regions logically belonging together. This approach however does not yield enough differentiated regions in order to constitute a usable graph (a graph of 5 vertices cannot topologically analyzed in any useful way). Thus the major traditional goal of image segmentation is incompatible with the goal of this work.

Nevertheless, segmentation does use techniques that could be adapted to generate hundreds or even thousands of smaller regions representing the different topological elements within an image – this is usually referred to as oversegmentation, yet it has already been used to generate finer grained partitions [54]. Depending on the algorithm, this can be accomplished by setting some region merging criteria to a higher threshold or adapting the rules for erecting watersheds.

**Supervised or Unsupervised Learning?** Because the final goal of most image processing techniques in medicine is to differentiate between healthy and pathological tissue, they belong to the group of problems known as classification problems, and are therefore supervised learning problems. However, the methods described above presuppose no anterior knowledge about the input images (or parts thereof) in order to group regions of pixels or features together, so the segmentation is done in an unsupervised fashion. This is certainly not the only possibility, as templates of individual features could be provided to the algorithm. Then again, the method would lose its generality, as different templates would be needed for different types of images. A possible solution to this problem is discussed later.

**What Information to Put into a Feature Vector?** Once a satisfying segmentation is obtained, some representative information has to be extracted from the individual regions in order to be stored as the feature vector of the resulting graph node. A whole phalanx of region properties can be chosen, and some will make more sense than others for a particular purpose. Aside from basic geometric information (centroid coordinates, length, or length-to-width ratio) [55] describes common features like histogram-based (mean grey values or grey level entropy distribution), pixel-co-occurrence related (angular moment, correlation, sum variance) as well as frequency-based (such as the wavelet) properties.

## 7 Future Challenges

**Computational Efficiency.** In comparison to extracting point cloud data from text documents, multimedia content such as images or video streams contain a very large amount of data (that might not necessarily contain much information).

As a simple example, in order to represent a 5 megapixel image as an adjacency graph, 25 billion entries would be necessary, while a more efficient representation as an adjacency list would still hold (depending on the neighborhood-definition) on the order of several dozen million list entries. This calls for efficient segmentation and clustering algorithms, as even quadratic runtime complexity would result in unacceptable computing times for interactive data mining in a real-world working environment. Possible solutions comprise downsizing images and the exclusive use of algorithms with near-linear ( $O(n * \log(n))$  being acceptable) runtime behaviour, as several graph-based algorithms like MST fortunately exhibit. Moreover, depending on the features selected to extract per output node, additional computation will be needed. While this may result in computing times acceptable for professionals depending on that particular information, it might not be to others, which calls for the inclusion of the professional end user into the data mining process.

**User Interaction Pipeline.** Although most algorithms discussed can produce results in a purely unsupervised fashion, in order to achieve excellent and relevant results, we propose designing an interactive data mining work flow. For example, a trained medical professional could identify regions-of-interest in an image which are then utilized by our algorithms to extract templates (feature vectors of those regions) for further use in future classification tasks. While most algorithms proposed today focus on very narrow fields of application (colon images, melanoma samples etc.), this would add to our software the flexibility to include per-user parameters into its machine learning process, solving the problem of what feature vectors to extract, thus significantly widening the applicability of our work.

**Visualizing  $n$ -Dimensional Point Clouds as Topological Landscapes.** A very promising research route has been opened by [56], [57], [58]: they utilize a landscape metaphor to images, which presents clusters and their nesting as hills whose height, width, and shape reflect cluster coherence, size, and stability. A second local analysis phase utilizes this global structural knowledge to select individual clusters, or point sets, for further, localized data analysis. The big advantage is that the focus on structural entities significantly reduces visual clutter in established geometric visualizations and permits a more efficient data analysis. This analysis complements the global topological perspective and enables the end user to study subspaces or geometric properties, such as shape. This is a very promising research route to follow.

## 8 Conclusion

Much further promising research routes are open for further exploration in the discovery of knowledge from natural images, however, the first question is how to preprocess the raw data as to get relevant data which is applicable for the use

of methods from geometry and topology. As this paper only describes methods to extract point cloud data from different weakly structured sources, once a point cloud (or graph) is extracted, it will have to be topologically analysed in order to produce workable results. The quality of those results will not only depend on the quality of the algorithms themselves, but to a large degree also on the quality of the input graphs they receive. In order to determine how well suited our graphs are for further computation, we will have to conduct those experiments, adapting our methods and parameters as needed.

## References

1. Holzinger, A., Dehmer, M., Jurisica, I.: Knowledge discovery and interactive data mining in bioinformatics state-of-the-art, future challenges and research directions. *BMC Bioinformatics* 15(suppl. 6), S1 (2014)
2. Edelsbrunner, H., Harer, J.L.: *Computational Topology: An Introduction*. American Mathematical Society, Providence (2010)
3. Memoli, F., Sapiro, G.: A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics* 5(3), 313–347 (2005)
4. Holzinger, A.: *Topological Data Mining in a Nutshell*. Springer, Heidelberg (2014) (in print)
5. Mmoli, F., Sapiro, G.: A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics* 5(3), 313–347 (2005)
6. Canutescu, A.A., Shelenkov, A.A., Dunbrack, R.L.: A graph-theory algorithm for rapid protein side-chain prediction. *Protein Science* 12(9), 2001–2014 (2003)
7. Zomorodian, A.: *Topology for computing*, vol. 16. Cambridge University Press, Cambridge (2005)
8. Vegter, G.: *Computational topology*, pp. 517–536. CRC Press, Inc., Boca Raton (2004)
9. Hatcher, A.: *Algebraic Topology*. Cambridge University Press, Cambridge (2002)
10. Cannon, J.W.: The recognition problem: What is a topological manifold? *Bulletin of the American Mathematical Society* 84(5), 832–866 (1978)
11. De Berg, M., Van Kreveld, M., Overmars, M., Schwarzkopf, O.C.: *Computational geometry*, 3rd edn. Springer, Heidelberg (2008)
12. Aurenhammer, F.: Voronoi diagrams - a survey of a fundamental geometric data structure. *Computing Surveys* 23(3), 345–405 (1991)
13. Axelsson, P.E.: Processing of laser scanner data - algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 54(2-3), 138–147 (1999)
14. Vosselman, G., Gorte, B.G., Sithole, G., Rabbani, T.: Recognising structure in laser scanner point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 46(8), 33–38 (2004)
15. Smisek, J., Jancosek, M., Pajdla, T.: *3D with Kinect*, pp. 3–25. Springer (2013)
16. Dal Mutto, C., Zanuttigh, P., Cortelazzo, G.M.: *Time-of-Flight Cameras and Microsoft Kinect*. Springer, Heidelberg (2012)
17. Khoshelham, K., Elberink, S.O.: Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors* 12(2), 1437–1454 (2012)

18. Kayama, H., Okamoto, K., Nishiguchi, S., Yamada, M., Kuroda, T., Aoyama, T.: Effect of a kinect-based exercise game on improving executive cognitive performance in community-dwelling elderly: Case control study. *Journal of Medical Internet Research* 16(2) (2014)
19. Gonzalez-Ortega, D., Diaz-Pernas, F.J., Martinez-Zarzuela, M., Anton-Rodriguez, M.: A kinect-based system for cognitive rehabilitation exercises monitoring. *Computer Methods and Programs in Biomedicine* 113(2), 620–631 (2014)
20. Holzinger, A., Dorner, S., Födinger, M., Valdez, A.C., Ziefle, M.: Chances of Increasing Youth Health Awareness through Mobile Wellness Applications. In: Leitner, G., Hitz, M., Holzinger, A. (eds.) *USAB 2010. LNCS*, vol. 6389, pp. 71–81. Springer, Heidelberg (2010)
21. Sitek, A., Huesman, R.H., Gullberg, G.T.: Tomographic reconstruction using an adaptive tetrahedral mesh defined by a point cloud. *IEEE Transactions on Medical Imaging* 25(9), 1172–1179 (2006)
22. Caramella, D., Bartolozzi, C.: 3D image processing: techniques and clinical applications (Medical Radiology / Diagnostic Imaging). Springer, London (2002)
23. Salton, G., Wong, A., Yang, C.S.: A vector space model for automatic indexing. *Commun. ACM* 18(11), 613–620 (1975)
24. Holzinger, A.: On Knowledge Discovery and Interactive Intelligent Visualization of Biomedical Data - Challenges in Human Computer Interaction & Biomedical Informatics. *INSTICC, Rome*, pp. 9–20 (2012)
25. Wagner, H., Dlotko, P., Mrozek, M.: Computational topology in text mining, pp. 68–78 (2012)
26. Argenziano, G., Soyer, H.P.: Dermoscopy of pigmented skin lesions—a valuable tool for early diagnosis of melanoma. *The Lancet Oncology* 2(7) (2001)
27. Eisemann, N., Waldmann, A., Katalinic, A.: Incidence of melanoma and changes in stage-specific incidence after implementation of skin cancer screening in Schleswig-Holstein. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 57, 77–83 (2014)
28. Argenziano, G., Giacomel, J., Zalaudek, I., Blum, A., Braun, R.P., Cabo, H., Halpern, A., Hofmann-Wellenhof, R., Malvehy, J., Marghoob, A.A., Menzies, S., Moscarella, E., Pellacani, G., Puig, S., Rabinovitz, H., Saida, T., Seidenari, S., Soyer, H.P., Stolz, W., Thomas, L., Kittler, H.: A Clinico-Dermoscopic Approach for Skin Cancer Screening. Recommendations Involving a Survey of the International Dermoscopy Society (2013)
29. Australia, M.I.: Dermoscopy (November 2013)
30. Ahlgrimm-Siess, V., Hofmann-Wellenhof, R., Cao, T., Oliviero, M., Scope, A., Rabinovitz, H.S.: Reflectance confocal microscopy in the daily practice. *Semin. Cutan. Med. Surg.* 28(3), 180–189 (2009)
31. Meijering, E., van Cappellen, G.: Biological image analysis primer (2006), booklet online available via [www.imagescience.org](http://www.imagescience.org)
32. Risser, J., Pressley, Z., Veledar, E., Washington, C., Chen, S.C.: The impact of total body photography on biopsy rate in patients from a pigmented lesion clinic. *Journal of the American Academy of Dermatology* 57(3), 428–434
33. Mikailov, A., Blechman, A.: Gigapixel photography for skin cancer surveillance: A novel alternative to total-body photography. *Cutis* 92(5), 241–243 (2013)
34. dos Santos, S., Brodlić, K.: Gaining understanding of multivariate and multidimensional data through visualization. *Computers & Graphics* 28(3), 311–325 (2004)
35. Emmert-Streib, F., de Matos Simoes, R., Glazko, G., McDade, S., Haibe-Kains, B., Holzinger, A., Dehmer, M., Campbell, F.: Functional and genetic analysis of the colon cancer network. *BMC Bioinformatics* 15(suppl. 6), S6 (2014)

36. Bramer, M.: Principles of data mining, 2nd edn. Springer, Heidelberg (2013)
37. Kropatsch, W., Burge, M., Glantz, R.: Graphs in Image Analysis, pp. 179–197. Springer, New York (2001)
38. Palmieri, G., Sarantopoulos, P., Barnhill, R., Cochran, A.: 4. Current Clinical Pathology. In: Molecular Pathology of Melanocytic Skin Cancer, pp. 59–74. Springer, New York (2014)
39. Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A., Huntley, A.: Segmentation of skin cancer images. *Image and Vision Computing* 17(1), 65–74 (1999)
40. Argenziano, G., Soyer, H.P., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, M., Cerroni, L., De Rosa, G., Ferrara, G., Hofmann-Wellenhof, R., Landthaler, M., Menzies, S.W., Pehamberger, H., Piccolo, D., Rabinovitz, H.S., Schiffner, R., Staibano, S., Stolz, W., Bartenjev, I., Blum, A., Braun, R., Cabo, H., Carli, P., De Giorgi, V., Fleming, M.G., Grichnik, J.M., Grin, C.M., Halpern, A.C., Johr, R., Katz, B., Kenet, R.O., Kittler, H., Kreusch, J., Malvey, J., Mazzocchetti, G., Oliviero, M., Özdemir, F., Peris, K., Perotti, R., Perusquia, A., Pizzichetta, M.A., Puig, S., Rao, B., Rubegni, P., Saida, T., Scalvenzi, M., Seidenari, S., Stanganelli, I., Tanaka, M., Westerhoff, K., Wolf, I.H., Braun-Falco, O., Kerl, H., Nishikawa, T., Wolff, K., Kopf, A.W.: Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet. *Journal of the American Academy of Dermatology* 48, 679–693 (2003)
41. Ferri, M., Stanganelli, I.: Size functions for the morphological analysis of melanocytic lesions. *International Journal of Biomedical Imaging* 2010, 621357 (2010)
42. Pizzichetta, M.A., Stanganelli, I., Bono, R., Soyer, H.P., Magi, S., Canzonieri, V., Lanzanova, G., Annessi, G., Massone, C., Cerroni, L., Talamini, R.: Dermoscopic features of difficult melanoma. *Dermatologic Surgery: Official Publication for American Society for Dermatologic Surgery* 33, 91–99 (2007)
43. Ballard, D.H.: Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition* 13(2), 111–122 (1981)
44. Ruppertshofen, H., Lorenz, C., Rose, G., Schramm, H.: Discriminative generalized hough transform for object localization in medical images. *International Journal of Computer Assisted Radiology and Surgery* 8(4), 593–606 (2013)
45. Tsai, A., Yezzi Jr., A., Willsky, A.S.: Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Transactions on Image Processing* 10(8), 1169–1186 (2001)
46. de Mauro, C., Diligenti, M., Gori, M., Maggini, M.: Similarity learning for graph-based image representations. *Pattern Recognition Letters* 24(8), 1115–1122 (2003)
47. Bianchini, M., Gori, M., Mazzoni, P., Sarti, L., Scarselli, F.: Face Localization with Recursive Neural Networks. In: Apolloni, B., Marinaro, M., Tagliaferri, R. (eds.) WIRN 2003. LNCS, vol. 2859, pp. 99–105. Springer, Heidelberg (2003)
48. Chen, C., Freedman, D.: Topology noise removal for curve and surface evolution. In: Menze, B., Langs, G., Tu, Z., Criminisi, A. (eds.) MICCAI 2010. LNCS, vol. 6533, pp. 31–42. Springer, Heidelberg (2011)
49. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(6), 583–598 (1991)
50. Meyer, F.: The steepest watershed: from graphs to images. arXiv preprint arXiv:1204.2134 (2012)
51. Sonka, M., Hlavac, V., Boyle, R.: Image processing, analysis, and machine vision, 3rd edn. Cengage Learning (2007)

52. Rogowska, J.: Overview and fundamentals of medical image segmentation, pp. 69–85. Academic Press, Inc. (2000)
53. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* 59(2), 167–181 (2004)
54. Lee, Y.J., Grauman, K.: Object-graphs for context-aware visual category discovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34(2), 346–358 (2012)
55. Wiltgen, M., Gerger, A.: Automatic identification of diagnostic significant regions in confocal laser scanning microscopy of melanocytic skin tumors. *Methods of Information in Medicine*, 14–25 (2008)
56. Oesterling, P., Heine, C., Janicke, H., Scheuermann, G.: Visual analysis of high dimensional point clouds using topological landscapes. In: North, S., Shen, H.W., Vanwijk, J.J. (eds.) *IEEE Pacific Visualization Symposium 2010*, pp. 113–120. IEEE (2010)
57. Oesterling, P., Heine, C., Janicke, H., Scheuermann, G., Heyer, G.: Visualization of high-dimensional point clouds using their density distribution’s topology. *IEEE Transactions on Visualization and Computer Graphics* 17(11), 1547–1559 (2011)
58. Oesterling, P., Heine, C., Weber, G.H., Scheuermann, G.: Visualizing nd point clouds as topological landscape profiles to guide local data analysis. *IEEE Transactions on Visualization and Computer Graphics* 19(3), 514–526 (2013)