Facet-based Image Processing and Representation

Ales V. Michtchenko
Moscow State University, Department of Applied mathematics and Computer Science
E-mail: ales.michtchenko@usa.net

Abstract

A fundamental approach to analysis and representation of an Image on a scale between Whole-Image-scale and Pixel-scale is still not well developed in Computer Science. On the other hand, it is clear, that successful manipulation with local image statistics, such as local histograms, local texture features, etc, may provide a powerful tool for Image Processing, Understanding and Multimedia. This paper describes one of the possible ways to construct such an approach, based on a notion of subimages (facets). It is shown, that this approach may provide a competitive results in such applications, as adaptive segmentation, boarder detection, stereoscopic vision and motion tracking. Paper is concentrated on analysis of such local image statistics, as facets' historgams.

This method allows, from one hand, to build a non-uniform multiresolution image subdivision and, on another hand, an approximation to semantical tree of a scene. The connection between them may be very useful for Multimedia applications.

Keywords: *Image Representation,* ***Image Processing****, Segmentation.*

1. FACET REPRERSENTATION OF AN IMAGE AND SEGMENTATION AS SORTING

Let us impose a square mesh (facet-mesh) with a side, equal to 2^N pixels on an image, being processed (see figure 1). This divides the image into a number of subimages. Each subimage in a mesh's cell contains 2^{2N} pixels. We will count a number of cell's pixels N_{pix}, having the same colour i (weight of this colour). Let us call this local histogram ($N_{pix}(i)$; $i = \overline{1, Ncol}$) a facet.

Note, that colours $i = \overline{1, Ncol}$ may correspond to either RGB or HS palette, depending on do we want to neglect light effects or not.

The main methodology of this section is to construct an object-scale colour distributions from local (facet-scale) colour distributions and therefore adaptively segment objects.

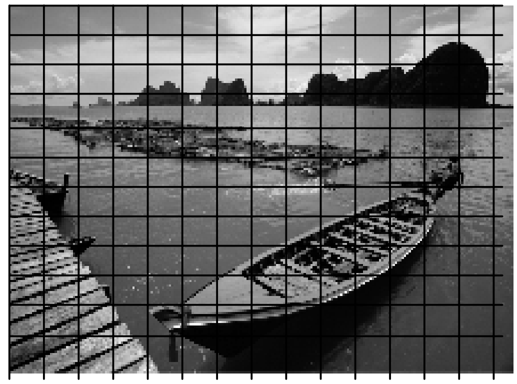

Figure 1:
Initial image with a facet-mesh imposed

Note, that instead of function $Npix(i)$; $i = \overline{1, Ncol}$, we may define a truncated function $Npix(i)$; $i = \overline{1, Nmaincol}$; $Nmaincol < Ncol$, limiting ourselves only with main colours of this facet.

We can treat a matrix from these facets as a blurred image, in which small objects are filtered out (see figure 2). Big objects in this matrix are represented by a set of facets. The set of facets, appeared "on the territory" of real object will be referred as a facet-object. Segmentation of real objects from the image will be guided by segmentation of facet-objects from facet-matrix. The latter is, actually, sorting facets into classes with similar main colours. After this is done, the object-scale colour distributions $Npix(i)$ may be received by averaging colour distributions of its facets.

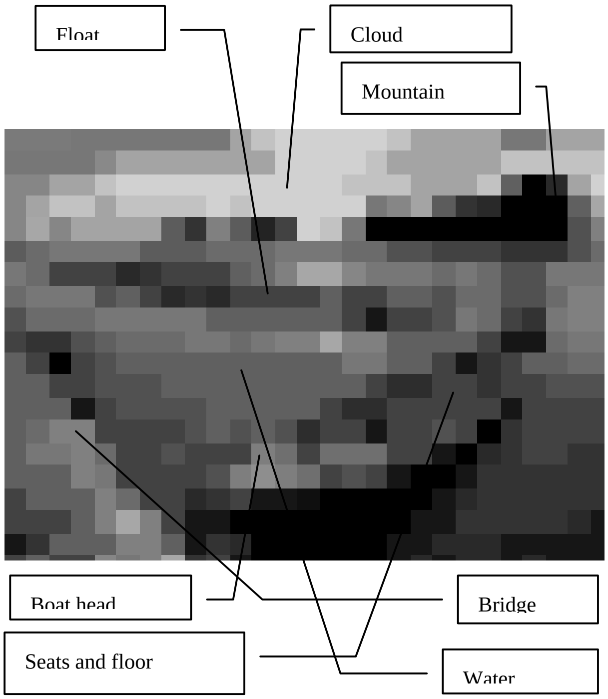

Figure 2:
Blurred image (facet size:16x16).
Some of detected facet-objects are shown.

Note, that classes of facets with similar main colours (future facet-objects) should be simply-connected. There are a sufficient number of methods, for image subdivision into simply-connected areas. However, in our case, the criterion for including a facet into a facet-object (colours similarity) is not straightforward. That is why we are proposing a special "Crystallisation" algorithm for segmentation of facet-objects from facet-matrix (see 1.2.).

1.1 Inner and boarder facets

We anticipate the problem of segmentation of facet-objects from facet-matrix by division of an initial set of facets into two subsets, namely: inner facets (belonging to the only one facet-object) and boarder facets (belonging simultaneously to two different facet-objects).

This type of segmentation continues the line of filtering small objects and details out. For example, a collection of small similarly coloured objects (fish jamb, tree crown, pile of papers) will be recognised as a single one. Moreover, transparent construction objects (like non-solid fences, trees without leaves, etc) will be segmented together with their background (see, for example a "Seats and floor" facet-object on a figure 2). In some cases, like segmenting complex construction as a single object, this coarsening makes sense. In other cases, this coarsening adds false facet-objects, like "middle" regions in places of complex occlusion/penetration of non-convex objects. But, in all cases, these effects are natural result of "blurring" the image and can be avoided only by making facets smaller (see section 2).

1.2 The algorithm of crystallisation

The algorithm of crystallisation of facet-objects crystallises (segments) objects in order of increasing amount of main colours in them and in order of decreasing the weight-difference between main and other colours. It works as follows:

First step (crystallising single-colour objects): Sorting all facets in the order of increasing difference $\Delta_1 = Npix(1) - Npix(2) - Npix(3) - \ldots - Npix(Ncol)$ between heavyset main colour $Npix(1)$ and other colours $Npix(i)$; $i = \overline{2, Ncol}$. The bigger the difference Δ_1, the bigger the probability, that a corresponding facet belongs to some single-colour object. Lowering an arbitrary threshold $\min \Delta_1$, allows more single-colour facet-objects to appear. After inner facets of single-colour facet-object are found, we are looking for boarder facets of this object. To find them, we are testing neighbours of all inner facets. If object's colour is among main colours of a neighbouring facet, this facet is treated as a boarder facet of this object.

Second step (crystallising two-colour objects): Analogously sorting all facets in the order of increasing difference $\Delta_2 = Npix(1) + Npix(2) - Npix(3) - \ldots - Npix(Ncol)$. Lowering an arbitrary threshold $\min \Delta_2$, allows more two-colours facet-objects to appear.

These steps are repeated up to $Nmaincol^{th}$ step. This order sets a priority: facets are rather treated as a boarder facet of object with less colours, than an inner facet of object with more colours. In particular, this treats T- and X-junctions as boarder between two heaviest colours.

1.3 Structure of facet-objects and quality of facet representation

Intuitively clear, that, the facet-based analysis of an image (analysing facet-object instead of real object) makes sense, if facets are small enough, so that at least one facet lies entirely in the real object (figure 3 a, c, d). Otherwise, the boarders of real object will lye in inner facets of a facet-object and no boarder facets will be detected (figure 3 b for unstructured object).

If most of inner facets border only with other inner facets or boarder facets, the facet object will be called structured (in objects 3a and 3c all inner facets satisfy this). More specifically, let us define a measure of structure of a facet-object as follows: Taking each inner facet, having non-inner neighbours and assigning a structure-value to it, equal to quantity of neighbours, belonging to the object (either boarder or inner facets). Taking the mean structure-value over all inner facets, having non-inner neighbours, we get the object's measure of structure. The facet-object will be called structured, if its measure of structure exceeds an appointed threshold.

The sum of object's measures of structure will be called a quality of a facet representation. This quality may be maximised over all possible thresholds ($\min \Delta_1$, $\min \Delta_2$, ..., $\min \Delta_{Nmaincol}$).

Finding an extremum of this function of $Nmaincol$ is a very computationally expensive procedure. It may be applied only if number of facets is relatively small. Then we will refine facet representation subdividing facet-mesh, we will use values of ($\min\Delta_1$, $\min\Delta_2$, ..., $\min\Delta_{Nmaincol}$), acquired for previous, coarse mesh (see section 2).

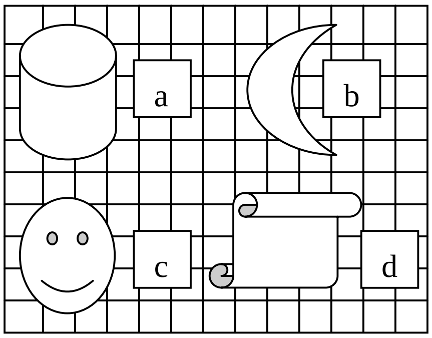

Figure 3:
Structured (a,c,d) and unstructured (b) objects.
Objects with a complex (c) and simple structure (a,d)

Apart from unstructured objects, it is possible to define a notion of object with a complex structure, as object, consisting from several smaller objects or object, containing smaller objects on its territory.

Image regions containing unstructured objects or objects with a complex structure are not adequately represented by this facet size. The next section describes, how these regions may subdivided further, leading to constructing non-uniform multiscale mesh and semantical tree of an image.

In the first case (unstructured objects) this subdivision may bring a structure to unstructured object (division into inner and boarder facets may appear). In the second case, complex-structure object may be divided into several simple-structure objects. For example, subdividing a facet-object "Seats and floor" on the figure 2, cause its division into a number of seats and floor-spaces between them. Subdividing a facet-object on figure 3c, cause finding on its "territory" a new objects, such as eyes and mouth.

2. HIERARCHICAL REPRESENTATION OF OBJECTS AND SCENE

Let us suppose that during mentioned above facet-based segmentation, we divided an image into a set of coarse objects. This set of objects consists from simple-structure objects, complex-structure objects and unstructured objects.

The next resolution level of facet segmentation is subdivision of facets, being component of complex-structure objects and unstructured objects.

This non-uniform partitioning of initial mesh ends then, at a certain level of resolution, all objects become simple-structured or then facet size reaches a certain minimum (we used 4x4 minimum cells, calling everything of more delicate structure a "motley" object).

The result is transforming an initial facet-mesh into non-uniform multiresolution mesh, which is tuned to be coarser in areas of simple geometry/texture and finer in areas of complicated geometry/texture.

Hierarchical representation of image [12] and objects ([11], [7]) was used for a wide range of vision applications. However, usually, a hierarchical representation of objects is based on pre-existing models and, therefore, is not automatic [7]. As for automatic algorithms for hierarchy construction, they are usually made to be applicable to a narrow range of specific tasks [12], [11]. In this subsection we are discussing, how a facet-based non-uniform multiresolution may result in automatic construction of hierarchical representation of objects and scene, useful for a general Vision and Recognition tasks.

Using the mentioned above multi-resolution facet-mesh, each object and a scene itself is represented by a tree. At the root of this tree there is a scene itself. The N-th level consists from objects found at the N-th resolution level. Leaves correspond to simple-structure objects or "motley" objects. Other nodes correspond to complex-structure objects or unstructured objects. A subtree, attached to complex-structure object, may represent smaller objects, found on its territory. A subtree, attached to unstructured object, represent the same object with better resolution.

3. MOTION ESTIMATION AND SEGMENTATION, OBJECT TRACKING

In this section we will briefly outline, how facet-based image representation could be applied to boarder detection, finding stereoscopic pairs and object tracking. A complete information on this subject can be found in [14].

As soon as a structured facet-object is acquired and it's main colours together with their weights are fixed, segmentation and analysis of a corresponding real (pixel-scale) object may be started.

Facet representation of an image facilitates different pixel-scale processing tasks, such as:

1) Boarder detection. Area of boarder search is narrowed to boarder facets. Besides that, boarder-edges are supposed to divide boarder-facets into areas of specific and non-specific colours, which are already defined for each facet.

2) Stereoscopic pairs finding. Facet representation allows to perform this task in a following order: finding pairs of similar facet-object, finding pairs of similar facets in them, and finally, pairs of edges in these facets.
Finding a stereoscopic pair can be based not only on epipolar range. The best candidate may be chosen using facet-based restrictions: size-similarity (it should consist from approximately

the same number of facets) and colour-similarity (its main colours should be approximately the same).

3) Motion tracking. The simplicity of finding stereoscopic pairs for facet-objects and for their facets means, that tracking dynamic objects may be performed similarly. Analysis of dynamic images may be based on finding the same objects and same facets in two subsequent video shots (the same way, as we did with two stereo images – [14]).

4. CONCLUSIONS AND EXPERIMENTAL RESULTS

Wide-scale testing of this method is not completed yet. However, some experiments with different indoor and outdoor images showed, that this method of image representation and processing can be applied both to constructing an approximation to semantical tree of an image and as a preprocessing stage for other image processing tasks (see section 3).

The method, being computationally expensive in itself, gives no performance increase if only one particular task is to be solved. However, if an image is to be used for different purposes, such as boarder detection, stereoscopic vision tasks, motion tracking, etc, an overall performance can be increased since facet-based representation gives a lot of information, useful for many image processing and vision tasks.

This multiscale hierarchical representation of objects and entire image, being processed, seems to be one of mechanisms used in biological vision, especially in image understanding, where at first, a coarse (blurred) picture is captured and, then, necessary small details are fixed by taking a peer look.

The results of tests showed, that this method has a big potential and could be adapted to different types of images processed and problems posed. This makes facet-based image processing an interesting new approach to Computer vision and Image Representation.

5. REFERENCES

[1] R.M.Haralick, L.G.Shapiro. *Computer and Robot Vision*. Addison-Wesley, Reading, Mass., 1993

[2] P.Besl, R.Jain. *Segmentation through variable-order surface fitting*. IEEE trans. Pattern Analysis and Machine Intelligence, Vol.10, N2, 1988, pp.167-192

[3] T.G.Fan, G.Medoni, R.Nevatia. *Segmented descriptions of 3D surfaces*. IEEE trans. Robotics and automation, Vol.3, N6, 1987, pp.527-538

[4] B.K.P.Horn. *Robot Vision*. Cambridge, MA, MIT Press, 1989

[5] G.Healey *Using colour for geometry-intensive segmentation*. J.Optical Society of America, June 1989, pp.920-937

[6] L.T.Maloney, B.A.Wandell *Colour constancy: A method for recovering surface spectral reflectance*. J.Optical Society of America, January 1986, pp.29-33

[7] G.J.Ettinger *Large hierarchical object recognition using libraries of parameterised model sub-parts.* In Proceedings of Computer Vision and Pattern Recognition, 1988, pp.32-41

[8] G.R.Cross, A.K.Jain. *Markov random field texture models*. IEEE trans. Pattern Analysis and Machine Intelligence, January 1983, pp.25-39

[9] J.Malik, P.Perona. *Preattentive texture discrimination with early vision mechanisms*. J.Optical Society of America, May 1990, pp.923-932

[10] V.S. Nalwa, T.O.Binford. Ondetecting edges. . IEEE trans. Pattern Analysis and Machine Intelligence, Vol.8, N6, 1986, pp.699-714

[11] L.M.Lifshitz, S.M.Pizer. *A multiresolution hierarchical approach to image segmentation based on intensity extrema*. IEEE trans. Pattern Analysis and Machine Intelligence, June 1990, pp.529-540

[12] Y.C.Chang, B.K.Shyn, J.S.Wang. *Region-based fractal image compression with quadtree segmentation*. In Proceedings of ICASSP-97, Munich, Germany, 1997.

[13] *From edges-to-regions approach and multiplayered image processing*, in Proceedings of "GraphiCon-99" pp.144-150, 1999

[14] A.Michtchenko. PhD thesis, Moscow State University, department of Computer Science, 2000

About the author

Ales V. Michtchenko is a PhD student of Moscow State University.

E-mail: **ales.michtchenko@usa.net**.