# Hierarchical Segmentation of Sparse Surface Data Using Energy-Minimization Approach

**Raid Al-Tahir**

Centre for Caribbean Land and Environmental Appraisal Research (CLEAR)
Department of Surveying and Land Information
The University of the West Indies, Trinidad and Tobago
rtahir@eng.uwi.tt

*Abstract*: The main objective for this research is to develop an algorithm that produces a dense representation of a surface from a sparse set of observations and facilitates preliminary labeling of discontinuities in the surface. The solution to these issues is of a great interest to the new trends and applications in digital photogrammetry, particularly for large-scale urban imagery.

This study adopts the approach of a concurrent interpolation of the surface and detection of its discontinuities by the weak membrane. The solution was achieved through developing a multigrid implementation of the Graduate Non-Convexity (GNC) algorithm. The conducted experiments proved that the developed method is adequate and applicable for dealing with large-scale images of urban areas as it was successful in producing a realistic surface representation and fulfilling other set criteria.

*Key words:* Surface Reconstruction, Discontinuioty Detection, Multigrid Regularization.

# 1 Introduction

Surface interpolation is a common and important task for several disciplines and applications in geosciences and engineering. This topic has regain research interest with the emergence of new trends and technologies in the collection of geo-spatial data such as digital photogrammetry and lidar. Such systems provide a large amount of data in the form of discrete points of three-dimensional coordinates.

Photogrammetry is a 3-dimensional coordinate measuring technique that uses mainly aerial photographs as the fundamental medium for measurements. The basic mode of operation is based on taking photographs from at least two different view points; light rays are then traced back from each photograph to points on the object and mathematically intersected to produce the 3-dimensional coordinates for these points [1]. The advances in computer technology, digital cameras, and the increasing resolution of recorded images have allowed more optimal utilization and efficiency, as well as developing new applications [2].

The tasks of surface interpolation and surface analysis play an important role in reconstructing the surface based on photogrammetric data. Properties of visual surfaces, such as breaklines and abrupt discontinuities in surface normals, must be made explicit. They have a critical part for the success of the earlier processes of image matching and stereopsis. These processes are usually executed over a hierarchy of image resolutions (image pyramid) [3]. On another aspect, the performance of any surface interpolation method will suffer greatly without incorporating such discontinuities. The matter becomes more critical particularly for large-scale urban imagery normally associated with higher percentage of occlusion, repeated patterns, and many surface discontinuities [4].

The large quantity of data generated by a photogrammetric system makes it a challenge to deduce 3D object descriptions from such data. Visual surface discontinuities are definite indications for boundaries of objects on the topographic surface [3]. These boundaries are crucial for furnishing appropriate surface representation, and providing the means for surface segmentation and object recognition [5]. Identifying spatial discontinuities is vital in many applications such as segmentation, optical flow, stereo, image reconstruction [6], and extracting high-level cartographic features (e.g., building representations) needed in cartographic analysis and visualization [3].

Despite the significance of surface interpolation and surface analysis, not all algorithmic and implementation aspects for solving the related subtasks are fully developed. There is no straightforward choice for an optimal surface interpolation method that would accommodate the needs. Thus, the objective for this study is to develop an algorithm that produces a dense representation for a surface from a set of sparse observations and, at the same time, facilitates the preliminary labeling of discontinuities in the surface.

# 2 Surface Interpolation and Discontinuity Detection

Surface interpolation may be a simple and well-understood routine for some applications, or a more complicated and critical issue for others. The reason behind this is the fact that the interpolation is a prediction of what is not known that would agree with the data to some extent and behaves reasonably between data points. Based on the density and the distribution of the data points as well as the computational principle, different methods provide different interpretations of the data, and thus, different representations of the surface may result [7].

A reasonable expectation for an adapted method for surface interpolation is to construct as realistic a surface representation as possible. In addition, it should preserve essential surface characteristics implied by the observations. A violation to this condition occurs when a smooth surface is interpolated over breaklines. Preferably, the adopted algorithm should refrain from introducing new characteristics to the surface, for example, creating new maxima or minima in the surface. It should also allow incorporating and utilizing information on discontinuities whenever such information becomes available [4].

The number of methods for interpolating data over a grid is large if one considers all different variations and solution methods [8]. Several methods were reviewed, tested and evaluated according to the aforementioned criteria. The conclusion was that interpolating a surface by membrane splines fulfills the criteria above, and provides explicit information for surface analysis [4]. Depth data, in this physical/mechanical rendering, are represented by a set of vertical pins scattered within the region; the height of an individual pin is related to the elevation of the point. Fitting a surface is then analogous to constraining a membrane (thin film) to pass over the tips of the pins. Such a model would generate minimal area surface that is continuous, but need not have continuous first partial derivatives [8].

On a different aspect of the issue, and since they assume a smooth and continuous surface, the simplistic and direct implementations of conventional methods for surface interpolation may not be adequate when encountering a breakline or discontinuity in the surface. The interpolated surface will exhibit an unrealistic behavior that is manifested as oscillation and overshooting. Hence, discontinuities must be detected at a certain phase of the interpolation process.

To address the issue of detecting discontinuities in the surface, several research studies have adopted the concept of a "line process", first introduced by [9]. A line process is a set of binary variables located at the lines that connect grid cells representing the surface. The purpose of a line process is to decouple adjacent cells if the values of these cells are different. However, a penalty should be paid when a breakline is introduced. Thus, a break line will only be introduced when paying the penalty (within a cost or energy function) is less expensive than not having the break line at all. Pre-setting the line process variable associated with specific cells can accommodate available information on their boundaries.

## 2.1 Energy Minimization Approach

Because some information is lost in the two-dimensional photographic recording and measuring processes, reconstructing the surface from a discrete set of data is an ill-posed problem [3]. Thus, regularization is normally used for obtaining the unique solution to the problem. It involves minimizing an energy function E that essentially includes two functionals. The first functional, *D(s)*, provides a measure of the closeness of the solution *(s)* to the available data. It is usually taken to be the sum of the square of the differences between interpolated values and the original data. The second functional, $S(s)$, measures the extent to which the solution conforms to the underlying assumption, usually expressed by the smoothness of the solution [8].

The minimization of $E$ is a trade-off between maintaining closeness to the original data and obtaining a smooth surface. The regularization parameter, $\lambda^2$, controls the influence of the functionals. A penalty function is added to the original energy function $E$. This function takes the form $P = \alpha l_i$, where $\alpha$ is the penalty and $l_i$ is the line process. This produces a combination of a continuous function for the surface and a discrete one for the lines:

$$E = D(s) + \lambda^2 S(s)(1 - l_i) + P. \qquad (1)$$

The energy functional in equation (1) prefers continuity in the surface, but allows occasional discontinuities if that make for a simpler overall description; a theme called "weak continuity constraints" [10]. This combination facilitates achieving surface reconstruction and discontinuity detection at the same time.

## 2.2 Graduated Nonconvexity Algorithm

The energy function $E$ is a non-convex function that has many local minima. In such a case, variational methods are used to guarantee obtaining the global minimum. Accordingly, the solution is achieved by means of embedding the functional $E$ into a one-parameter family of functionals $F^{(p)}$. The parameter $p$ represents a sequence of numbers ranging from one to zero. The function $F^1$ is a crude, but convex, approximation to the non-convex function. However, as $p$ goes to zero, $F^p$ becomes closer to the original non-convex one [11].

To obtain the function sequence $F^{(p)}$, the approach of Graduate Non-Convexity (GNC) algorithm developed by [10] is adopted. Accordingly, the line process $P$ is first merged with the interpolation function $S$ to create the neighbor interaction function ($g_{\alpha\lambda}$) that is not continuous (Figure 1). The modified configuration is then solved by the graduated non-convexity algorithm, where the non-convex function $E$ is gradually approximated by a convex one, $F$, through a family of $p$ intermediate functions. The final solution represents a convex approximation that encompasses the original function and has only one global minimum. The neighbor interaction function is also modified accordingly and approximated with a set of splines, as shown in Figure 1, denoted as $g_{\alpha\lambda}^{(p)}(t)$, and expressed as

$$g_{\alpha\lambda}^{(p)}(t) = \begin{cases} \lambda^2 t^2 & \text{if } |t| < q \\ \alpha - c(|t| - r)^2/2 & \text{if } q \le |t| < r \\ \alpha & \text{if } |t| \ge r \end{cases} \qquad (2)$$

where $c = 0.25/p$, $r = \alpha(2/c + 1/\lambda^2)$, and $q = \alpha/(\lambda^2 r)$.

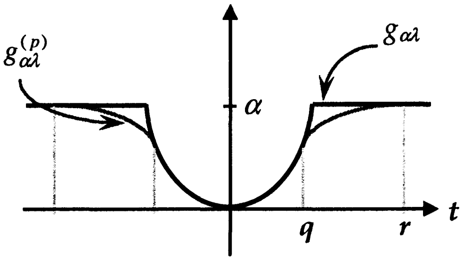


Figure 1. Approximation of the neighbor interaction function $g_{\alpha\lambda}$.

Numerical solution of the function sequence $F^{(p)}$ is achieved by implementing a relaxation algorithm, all terms of which are functions of the parameters $\lambda$ and $\alpha$. Beside their original connotation, the first parameter defines the distance for the interaction between discontinuities, beyond which two discontinuities do not interfere with each other. The second parameter ($\alpha$) is "immunity to noise" measure that prevents generating spurious discontinuities when $\alpha > 2\sigma^2$, where $\sigma^2$ is the variance of the mean noise [10].

There are other aspects of performance that depend on these parameters in discontinuity detection. The first one is contrast sensitivity threshold ($h_0 = \sqrt{2\alpha / \lambda}$), which determines the minimum contrast for detecting an isolated step edge. The other aspect is the gradient limit ($g_l = h_o / 2\lambda$) that would identify a ramp as a step when its gradient exceeds $g_l$. These different aspects of performance in discontinuity detection provide guidelines for defining the values of $\lambda$ and $\alpha$. Based on the broad expectations of the surface, one may choose specific values for the parameters considering, e.g., the desired height of a step to be detected ($h_o$) [4].

# 3 Multigrid Framework Development

The Graduated Non-Convexity solution for the weak membrane has the virtue of simultaneous execution of surface interpolation and discontinuity detection. However, sparse data set poses an intricacy at deciding on the exact location of a discontinuity when it is required. This ambiguity arises from the fact that inserting a breakline anywhere between data points that are several grid nodes apart will yield the same low energy. Because of that, it was recommended that the sparse data must first be converted to a dense one using a continuous membrane [10], [4]. Thus, carrying out surface interpolation and analysis tasks would involve two main sub-tasks. The first one is surface densification that produces a preliminary dense representation of the data. The second sub-task is the actual combined process of interpolation and discontinuity detection by the weak membrane. Both sub-tasks involve applying a relaxation method.

## 3.1 Relaxation and Multigrid Methods

A numerical solution by relaxation techniques (e.g., Gauss-Sidel, and successive over-relaxation methods) would approximate the differential equation by finite differences (or finite elements) [12]. A grid representation of the surface is then obtained by computing its value at each node of the grid using a weighted sum of the values of the neighboring nodes. The solution is acquired through an iterative fashion, where each new iteration improves the attained surface representation.

Standard relaxation methods are inherently inefficient in propagating information over large region of no data [13]. The method of multigrid provides the remedy for dealing with low-frequency components, and speeding up the convergence. It provides two strategies; the coarse grid correction and the nested iteration. The principle of the coarse grid correction scheme is to transfer (restrict) the solution to a coarser grid when relaxation begins to stall (due to reaching smooth error modes). The low frequencies on a fine grid look less smooth (i.e., become higher ones) on a coarser grid. The process is then repeated with the possibility of transferring to yet coarser grid [13].

The other strategy in multigrid computations is the nested iteration scheme, which prescribes that the initial approximation for a fine grid is obtained by interpolating (prolonging) the solution of the coarser grid, instead of starting with any arbitrary approximation. The solution of the coarser grid, in turn, is found by the same process from even coarser grids. The solution of the coarsest level is obtained by exact, or more rigorous, solution of the problem. With such an approach, there is a smaller number of unknowns to update, smaller number of relaxation sweeps, and, thus, faster rate of convergence. Nested iteration is then a recursive application of the correction scheme [13].

## 3.2 Multigrid GNC for the Weak Membrane

This study opted for implementing the GNC algorithm in a multigrid scheme. As the result of pursuing the multigrid approach, gaps representing the low frequencies in the surface would be filled, and better and faster surface representation would be produced. Implementing the tasks of surface densification and the weak membrane within a multigrid scheme will therefore improve their performance.

The multigrid implementation requires defining two different structures; one is related to image pyramid and another for surface grid resolution. For every level in the image pyramid, there is a surface representation of the same resolution in the object space. However, the surface is further sampled into grids of coarser resolution to satisfy the algorithmic requirements for providing an initial approximation for the surface. This proposition is demonstrate by Figure 2.

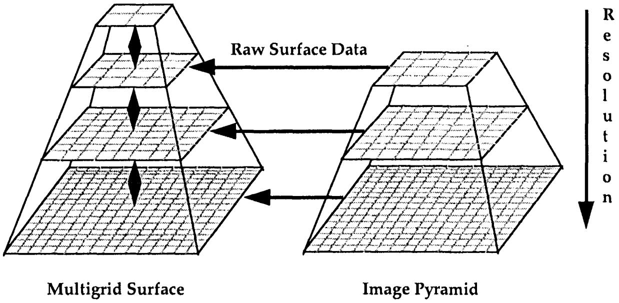


Figure 2. The relation between image pyramid and multigrid surface.

The Multigrid GNC, thus, will proceed through several phases. First of all, the original data must be preprocessed and sampled according to the definitions of the grid at each level. Secondly, a complete cycle of the multigrid membrane would be performed to obtain a dens surface at a specific level of surface representation using a prolonged solution from the previous level as well as the corresponding data. At the end of such cycle, the GNC routine would be activated as the third phase. The results of performing the GNC algorithm are a surface representation and a set of detected discontinuities at the specific level of image resolution. These results would be used to provide an approximation for the new cycle at the next finer image resolution in the pyramid.

# 4. Test Data and Experiments

The test data is related to a multi-story building with surrounding walkways and landscape that is visible in two consecutive aerial photographs (stereo pair). The left part of Figure 3 depicts the building in a section of one of the aerial images. The photographs were scanned at two image resolutions. The coarse resolution provides an image of a pixel size equivalent to 2 meters on the ground, while the finer resolution corresponds to 1-meter resolution.

The image-matching phase in the photogrammetric reconstruction of visible surfaces was conducted using the two image resolutions. The outcome was two sets of points, each of which represents three-dimensional points in successfully matched edge segments in the corresponding image resolution. The center and right parts of Figure 3 represent the data points obtained at the 2-meter and 1-meter image resolutions.

In order to proceed with the study, the first set (2-meter resolution) had to be sampled into three levels of cell size equals 2, 4, and 8 meters. The criterion used to determine the number of grids was the percentage of the nodes with data versus nodes without data (50% or more). Such a value was set empirically in order to promote small number of consecutive nodes of no data in a grid.

Figure 3. A section of a digital image showing the test site (left), the matched edges at 2-meter (center), and 1-meter (right) resolution.

The three-level sampled data set was then used in a full cycle of the multigrid GNC methodology (i.e., a combined implementation of the multigrid membrane followed by the weak membrane). Figure 4 represents the resulting surface in the final level after a complete multigrid cycle, as well as the detected breaklines at this level.

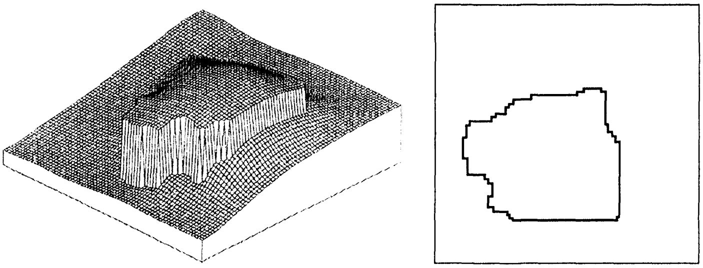

Figure 4. Surface representation for the 2-meter data resolution produced by multigrid GNC (left), and the detected discontinuities (right).

The last experiment attempted to apply the developed algorithm on multi-resolution data. Both data sets (corresponding to matched points at 2-metre and 1-meter image resolutions) were resampled to the required number of levels down to 8-meter resolution. Figure 5 shows the surface as interpolated by the multigrid implementation of the weak membrane, and the detected discontinuities.

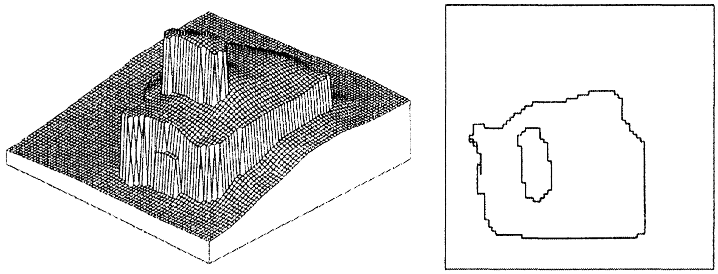

Figure 5. Surface representation for the 1-meter data resolution produced by multigrid GNC (left), and the detected discontinuities (right).

# 5 Concluding Remarks

The main contribution of this study was the development of a multigrid implementation of the Graduated Non-Convexity (GNC) algorithm for the weak membrane. The algorithm presented by this study offers the solution for the problems associated with the original GNC algorithm in dealing with sparse set of data. The conducted experiments proved that this algorithm is adequate and applicable for dealing with large-scale images of urban areas.

The multigrid formulation of the GNC offers an additional benefit as it lends itself to a more objective definition of the required GNC control parameters. The value of $\lambda$ (the scale) can be determined in conjunction with the level of surface grid. At the lowest resolution of the grid, it is chosen a bit larger than one (e.g., 1.1) to keep the smoothing and breaking of the surface at the minimum. However, $\lambda$ would be increased by one each time the algorithm progresses to the next higher level. As with respect to the contrast threshold ($h_o$), this study found that GNC performs better when $h_o$ is close to twice the mean deviation from the mean of the surface. This way ho is related to the surface data and its variations.

Hierarchy is an intrinsic concept of visual surface reconstruction. Thus, multigrid scheme would provide the means for the communication between the different levels of the image pyramid and among the different phases of surface reconstruction (e.g., image matching and object recognition). In addition to dense interpolated surface, the multigrid representation of the line process serves to communicate information on detected breaklines to these components.

Besides influencing the matching process at the next level of image resolution, there are several ways to benefit from the detected discontinuities. They can be compared or incorporated with other indications. One of which is the set of building boundaries in the object-space like those inferred from lidar observations [5] or from other evidences in the aerial photos after registering and integrating both data in a fashion similar to [14]. Another way is by studying the scale-space behavior of discontinuities detected in the surface at the end of each cycle [15].

## References


[1] Mikhail E., Bethel J. and McGlone C. 2001. *Modern Photogrammetry*, John Wiley, USA.

[2] Keating T., Garland P. and Dörstel C. 2003. 'Photogrammetry Goes Digital,. *GIS Development.*

[3] Schenk T. 1999. *Digital Photogrammetry*, Volume 1, TerraScience, OH, USA.

[4] Al-Tahir R. 1996. *Interpolation and Analysis in Hierarchical Surface Reconstruction,* Report 435, Department of Geodetic Science and Surveying, The Ohio State University, OH.

[5] Al-Tahir R. 2003. 'Segmentation of Lidar surface data using energy-minimization approach'. *Conradi Research Review* 2(2), pp. 76-85.

[6] Hewer G., Kenney C. and Manjunath B.S. 1998. 'Variational image segmentation using boundary functions'. *IEEE Transactions on Image Processing* 7(9), pp. 1269-1282.

[7] Watson D. 1992. *Contouring: A Guide to the Analysis and Display of Spatial Data.* Pergamon Press.

[8] Wolberg G. 1990. *Digital Image Warping*, IEEE Computer Society Press, CA.

[9] Geman S. and Geman D. 1984. 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6(6), pp. 721-74.

[10] Blake A. and Zisserman A. 1987. *Visual Reconstruction*, MIT Press, MA.

[11] Nielsen M. 1997. 'Graduated Nonconvexity by Functional Focusing'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(5), pp. 521-525.

[12] Nakamura S. 1991. *Applied Numerical Methods with Software*, Prentice Hall, NJ.

[13] Briggs W., Henson V. and McCormick S. 2000. *A Multigrid Tutorial*, 2nd edition, Society for Industrial & Applied Mathematics, Pennsylvania.

[14] Schenk T. and Csathó B. 2002. 'Fusion of Lidar data and aerial imagery for a more complete surface description'. *Proceedings of ISPRS Symposium Photogrammetric Computer Vision* (Austria) IAPRS 34(3A), pp 310-317.

[15] Witkin A., Terzopoulos D. and Kass M. 1987. 'Signal matching through scale space', *International Journal of Computer Vision* 1, pp. 133-144.