Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı, S. 509-521, Eylül 2020
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue, pp. 509-521, September 2020
Copyright © 2020 EJOSAT
Research Article

Kenar Bulma Algoritmaları ile 2 Boyuttan 3 Boyuta Dönüşüm

Cem Su Karaca [1*], Kemal Tütüncü

[1*] Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Savunma Teknolojileri Bölümü, Konya, Türkiye (ORCID: 0000-0002-5605-9074)
[2]Selçuk Üniversitesi, Teknoloji Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Konya, Türkiye (ORCID: 0000-0002-3005-374X)

(1st International Conference on Computer, Electrical and Electronic Sciences ICCEES 2020 – 8-10 October 2020)

(**DOI:** 10.31590/ejosat.828404)

ATIF/REFERENCE: Cem Su KARACA. & Kemal TÜTÜNCÜ. (2020). Conversion from 2 Dimension to 3 Dimension with Edge Detection Algorithms. *European Journal of Science and Technology*, (Special Issue), 509-521.

Abstract

In this thesis, a 3D content will be created from a 2D content. In these techniques, there is only one image and the aim is to create the second image using this image information. The most commonly used 2D to 3D conversion method is based on the depth fact, and in this way DIBR (Depth image-based rendering) is also the demand of the second image.
In this thesis, Matlab software will be used as the material that can provide data. 3 resizing, resizing picture, Difference Variable, Linear Depth Map Perpecific Direction, LDH start and end and should be done if desired.
Segmented Image, Depth Map, Left image, Right image, 3D image that we can obtain as output. Thus, the system can perform minimum input 3 sizing. In addition, this method can be applied in videos, although it is not synchronous.
As a result of the researches and examinations, it has been seen that there are numerous methods in the process of converting from two dimensions to three dimensions. It has been determined in the investigations that there are weak and strong points of each method.
The method was examined and it was seen that it was divided into two main topics, automatic and semi-automatic. Most of the transmission method today has focused on automatic depth extraction, but the decline in user usage has made it difficult to control information results. It has been observed that some errors could not be solved. It has been determined that correct returns cannot be taken regarding the depth information of the objects and some processes may be required. In order to avoid these disadvantages, the user usage used in semi-automatic usage is considered to be an important advantage.
It has been examined and seen in the experimental results of those who propose the method that the methods included in the methods we propose can obtain stereoscopic images that can be sorted in the correct order. In the next process, guiding critical information was assimilated in the software applications of these methods and the methods were examined in theoretical detail.

Keywords: Image, 3D (3-Dimensional), Edge Information Method, Graph Cuts and Random Walks, Depth

Conversion from 2 Dimension to 3 Dimension with Edge Detection Algorithms

Öz

Bu tez çalışmasında 2B bir içerikten 3B bir içerik oluşturma teknikleri olacaktır. Bu tekniklerde tek bir görüntü mevcuttur ve amaç bu görüntü bilgisini kullanarak ikinci görüntüyü oluşturmaktır. En sık kullanılan 2B'den 3B'ye çevrim yöntemi derinlik haritasının oluşturulması ve bu şekilde DIBR (Depth image based rendering) kullanılarak ikinci görüntünün oluşturulmasıdır.
Bu tez çalışmasın da veri setleri oluşturmak amacıyla materyal olarak Matlab yazılımı kullanılacaktır. Kullanılan bu sistemin girdileri 3 boyutlandırılması istenen resim, Fark Değişkeni,,Lineer Derinlik Haritası Perpesktif Yönü, LDH başlangıç ve son değerleri ve isteniyorsa obje bulunmasıdır.
Çıktı olarak elde edebileceğimiz Bölütlenmiş Görüntü, Derinlik Haritası, Sol görüntü, Sağ görüntü, 3 Boyutlu görüntüdür. Dolayısıyla sistem minumum girdi kullanarak 3 boyutlandırma işlemi yapabilmektedir.
Ayrıca , bu yöntem eş zamanlı olmamakla beraber videolar içinde uygulanabilir. Yapılan araştırmalar ve incelemeler sonucunda iki boyuttan üç boyuta dönüştürme işleminde sayısız yöntemin var olduğu görülmüştür.Her yöntemin zayıf ve güçlü olduğu noktaların

mevcut olduğu yapılan incelemelerde tespit edilmiştir.Görüntünün karakteristik özelliklerinin, kullanılacak dönüştürme yönteminin seçiminde doğrudan etkili olduğu anlaşılmıştır.
Dönüştürme yöntemleri incelenerek otomatik ve yarı otomatik olarak iki ana başlığa ayrıldığı görülmüştür. Günümüzde dönüştürme yöntemlerinin birçoğu otomatik derinlik çıkarımına odaklanmış, ancak kullanıcı müdahalesinin azalması, dönüştürme sonuçlarını kontrol etmeyi zorlaştırmıştır. Bazı hatalar kolayca çözülemediği görülmüştür.Nesnelerin derinlik bilgileriyle ilgili doğru geri dönüşler alınamayabileceği ve bazı ek süreçler gerekebileceği tespit edilmiştir. Bu dezavantajları bertaraf etmek için yarı-otomatik yaklaşımlarda kullanılan kullanıcı girişi önemli avantaj sağladığı kanısına varılmıştır.
Önerdiğimiz yöntemlerin içerdiği süreçlerin doğru bir şekilde uygulanması halinde kaliteli stereoskopik görüntüler elde edilebileceği, yöntemi önerenlerin deneysel sonuçlarında incelenmiş ve görülmüştür. Bundan sonraki süreçte bu yöntemlerin yazılım uygulamalarında yol gösterici olacak kritik bilgiler özümsenmiş ve yöntemler teorik açıdan ayrıntılı bir şekilde irdelenmiştir.

Anahtar Kelimeler: Görüntü, 3B(3 Boyutlu), Kenar Bilgisi Yöntemi, Grafik Kesimleri ve Rastgele Adımlar,Derinlik

1. Introduction

3D-three dimensional image technology has evolved and continues to evolve to ensure that images in two-dimensional planes from the past to today are perceived by humans in three dimensions as a result of various processes. The question that inspired this development process has been: How can the human eye perceive its environment in three dimensions?

In order to reach the answer to this question, the roadmap revealed the need to create the depth information needed to detect the two-dimensional image in three dimensions. It has been considered that if this depth information is perceived by the human eye, it will be possible to perceive two-dimensional images in three dimensions.

Man has two eyes and there is a certain distance between these two eyes (50-75 mm). So each eye sees its surroundings from a little different angles. Two different eyes detect two different images. While these two perceived images are very similar, they are not exactly the same. Balancing this small difference between two images requires a lot of geometric processing power. It is the human brain that performs this procedure. As a result of the perception of these two different images in the eye and the unification of the brain, the human environment can be seen in three dimensions.

This study will talk about various methods of obtaining a three-dimensional image and focus on techniques for converting a two-dimensional image to a three-dimensional image, and step-by-step analysis of depth map creation using edge information and depth map creation using the Chart Cuts and Random Walks method.

2. 2D to 3D Conversion

The part of creating 3D content that will be considered in this study is the techniques for creating 3D content from a 2D content. There is only one image in these techniques, and the goal is to create the second image using this image information. This is not a simple process. There are many research and development methods available in this regard. The most commonly used 2D-to-3D conversion method is the creation of a depth map, and in this way, the creation of the second image using DIBR (Depth image based rendering).

To create a three-dimensional image, you must first have depth information in each pixel. The picture created by these depth values is called a depth map. The depth map usually consists of 8-bit (256-digit) values, which are shown as a gray-toned image, 0 black and 255 white. In short, the depth map is a gray-scale image that shows the depth of pixels in an image. In general, those close to pixels are assigned high values on a gray scale, while the value assigned on a gray scale shrinks as objects move away. When looking at the depth map in approaches to 2D to 3D cycle, DIBR is used to create the second image.

Depth Image Based Rendering (the process of using a depth map throughout the image to create synthesized images of the scene. This is not an easy process and many problems can be encountered during its operation. The most important problem is visual clogging (optical occlusion). This event is defined as the pixel in one image not being in the other image or hiding it. In this case, the clogged information needs to be remanufactured. This process does not focus on blockages, it uses easier methods. The use of the depth map produced to create a second image can cause other problems, such as the fact that not all pixels are in the same image and some uneable pixels appear in the other image, in this way gaps appear, and so on.

Methods used to convert from 2D to 3D are generally examined under two headings.

2.1 Automatic Conversion Methods

Automatic 2D to 3D translation methods are based on different techniques based on depth knowledge in a 2D image. Depth information in the image is needed to create the depth map.

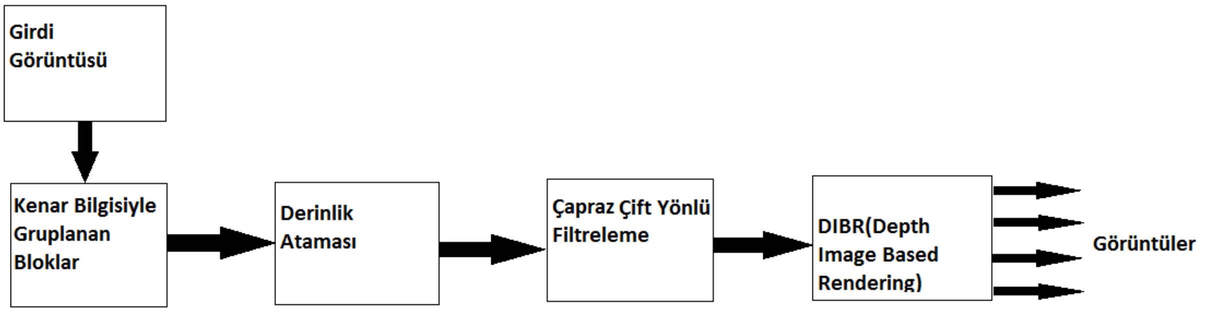
Some popular cycle methods are Depth from motion, depth from edge information, Color based depth extraction, and depth from geometry.

The current section will look at the Conversion method using Edge Information from the automatic conversion methods.

2.1.1 Converting Using Edge Information

2.1.1.1 Recommended System

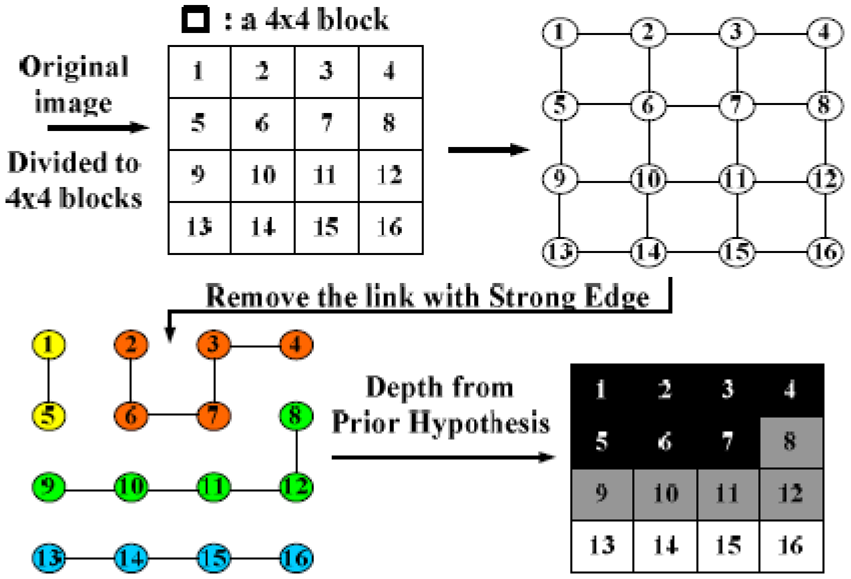
The process of located depth from a 2D image is a poorly positioned problem. But some depth clues and features can help us get to depth information. This method using edge information to provide conversion. The general concept of this method is this: The edge information of the image has a high potential to remove the edge information of the depth map. The block diagram of the proposed conversion system is given below.


Girdi Görüntüsü
Kenar Bilgisiyle Gruplanan Bloklar
Derinlik Ataması
Çapraz Çift Yönlü Filtreleme
DIBR(Depth Image Based Rendering)
Görüntüler

First, blocks by algorithm are used to reduce computational clutter. As shown in the way, the grouping stream uses edge information. An example is a 4x4 chart. Each node is a 4x4 pixel block. Each node has four connections. The value of each connection is the absolute difference between the averages of neighboring blocks.

$$\text{Diff (a,b)} = \left|Mean(a) - Mean(b)\right|$$

a and b represent two neighboring blocks. The Minimum Spanning Tree (MST) is created first. Stronger edges are removed to create a block group. Depending on the group numbers previously defined, edge links that are stronger for creating multiple grouped regions are removed. The reason the MST algorithm is used is because of the use of the MST algorithm. to define the harmony between color differences and blocks. The MST algorithm gives better results in spatial locations to maintain connections.

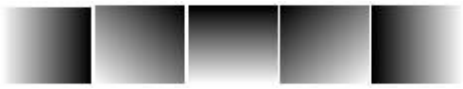


2.1.1.2 Depth from The Priority Hypothesis

The depth information corresponding to the depth (x,y) is generated from the corresponding position. The initial depth gradient hypothesis is shown below. Detailed depth assignment for a given block group is done as:

$$Depth(R)=128+255*\Big(\sum_{pixel(x,y)\in R}\Big(W_{rl}\Big(\frac{x-width/2}{width}\Big)+W_{ud}\Big(\frac{y-height/2}{height}\Big)\Big)/pixel\_num(R)\Big)$$

$$\left|W_{rl}\right|+\left|W_{ud}\right|=1$$

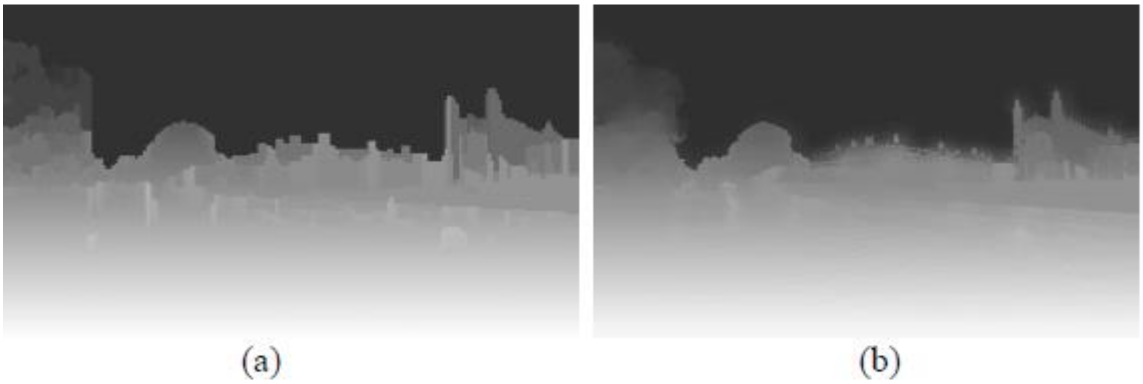


Integral Depth Gradient Map

This equation shows the center of gravity of the assigned depth value. W_{rl} and W_{ud} values set the depth gradient direction. The alignion of the default depth is obtained from the geometric perspective analysis of the image.

2.1.1.3 Dusk Filtering and Depth Map Based Image Scrolling (DIBR)

The cross-dui filter is applied to soften the depth of object boundaries. The filtered depth map has a more useful view quality. Because double-sided filtering produces a smooth depth map within smooth space with similar pixel values and protects the edge boundaries of objects. The filtered depth map is then used for three-dimensional imaging. Here's an example of double-sided filtering.



a) Depth assigned block image b)Double-sided filtered image

2.2 Semi-Automatic Conversion

Semi-automatic conversion from 2 dimensions to 3 dimensions; the conversion process with input from the user. There are several ways to implement user input into the system. However, the most commonly used are the user-provided marks. The marks correspond to the depth information in the picture. The main idea is to give hints to the algorithm to determine what depth is assigned to the pixels. Thus, a depth map can be created thanks to user-defined inputs.

The reason it is called semi-automatic is because there are some user interactions in the cycle process. These interactions are not output editing can be a user invocation process for priority information or more information in the algorithm. Semi-automated systems have recently received more attention than automated systems. This is because semi-automated systems do better than automated systems when used in large-scale productions. Semi-automated systems make concessions for higher quality in real-time applications. This is a much better job compared to all other conversion methods.

The current section will look at the Chart Segments and Random Steps method from the semi-automatic conversion methods.

2.2.1 Chart Cuts and Random Walks

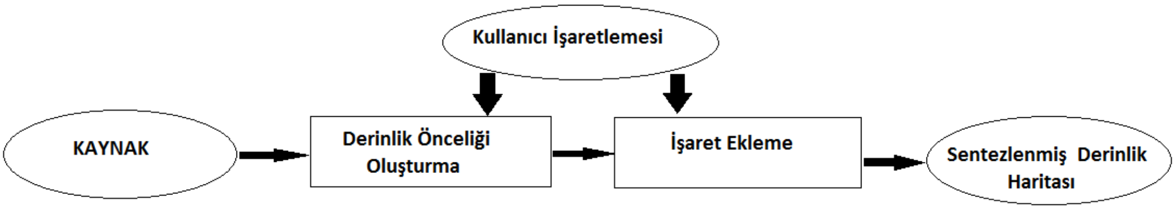
Today, many of the conversion methods are focused on automatic depth insi that is. However, reduced user intervention can make conversion results difficult to control. Some errors may not be easily resolved. Correct feedback on objects' depth information may not be possible, and some additional processes may be required. User input used in semi-automated approaches to eliminate these disadvantages provides a significant advantage.

For a single image or movie frame, the user simply marks objects and regions in light and dark colors, whichever is closer or farer to the camera. One question that often arises here is whether the marks applied by the user will occur in the actual scene perception. Depth marking doesn't have to be absolutely accurate. It just has to be perceptically consistent. User markup will meet the criteria in the method. As a result, the depths for the rest of the pixels will be predicted using this information. The semi-automatic conversion process allows the user to correct depth errors on the surface and eventually gives them a faster and more accurate conversion opportunity. It also offers a more cost-effective solution.

There are some common methods in this area. Specifically, the depths in a video series in the work done by Gutmann are resolved only by marking movements in the first and last film frames. The first step is to determine the detection points in all movie frames. Roughly depth estimation is made for the first and last film frame, and support vector machine (SVM) classifiers are run separately for both film frames. Each SVM is applied using scale-immutable property transformation (SIFT) for a unique depth classification in rough lycing. SIFT points are determined throughout the entire video and placed in SVM classifiers to determine detection points. The "one vs all scheme" is used by selecting the strongest similarity in all classifications for each SIFT point. The point at which this exceeds the high trust threshold is the point of determination, and the depth is assigned to the SVM classifier set for the specific depth. An energy function is minimized by the smallest squares to analyze the rest of the depths. This function allows you to use the is a combination of location and time-related concepts, color information, detection points, and user interventions. This minimizing process is done by the system of converting the problem into the sparse linear system of equations and solved directly. A similar method was developed by Wang. In this method, the problems of gaps (Holes) and clogging (Optical occlusion) are solved quite well. Together, they are simply caught near areas that do not have clogging. But the problem here is that this process requires a lot of processing load. It does not require calculating only SIFT points and SVM classifiers. But this requires solving large sequences of linear equations and high memory values.

2.2.1.1 Methodology

This method is user-defined at the beginning of depth eseses. Marks user objects and regions as close or farther. The method allows the user to mark with monochromatic densities, such as a color palette that changes from dark to light. The flow diagram is given in the following way.



2.2.1.2 Graphic Representation

To provide the highest level of flexibility in this method, images and videos are processed in the same way as N-linked graphics. From this point of view, a graphic is thought to be a mesh of nodes, where each pixel is a node. Pixels or nodes have interconnected edges. The edges are assigned to numeric values and show how similar the linked pixels are. An N-linked graphic is an N-edged representation in which each pixel is bound to another pixel. As shown in the example in the example, a 4x4 image is marked with different colors and three depths.

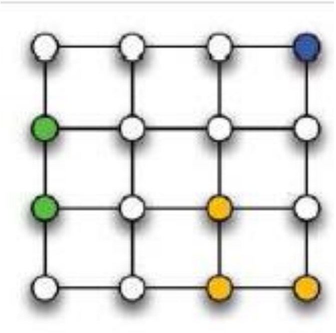


Figure 2.4x4 marked sample image

Colored nodes show user-marked pixels.

Each node is a pixel that is linked to its horizontal and vertical neighbors. So the example in the way is a 4x4 chart.

2.2.1.3 Transformation Structure for Images

In this method, graphical representation was used successfully in image segmentation, especially with 'Chart Separations and Random Steps'. This method used a combination of 'Chart Separations and Random Steps' to find an optimal depth markup in source images.'Chart Separations' is used for the initial depth map, which is taken as a priority depth (Depth Prior). This serves as an addition to the information channel in the 'Random Steps' algorithm.Now this method will look at how to create a depth map for each build individually.

2.2.1.4 Random Steps for Images

'Random Walks is an optimization program that covers all nodes that are not marked in charts, finds the probability of Random Steps, and starts at some signs. Steps are trended by edge weights. Thus, it is more likely to control similar pixels than different ones. The purpose of image segmentation is to classify each pixel in an image according to the possible label K to which it belongs. Random Steps determines the probability of the marks to which each pixel belongs. Signs or labels are the classification of high probability pixels. This is carried out by dissing the linear system based on laplacian matrix.

If vi shows the first pixel in an image, its size is RxC. Thus, the size of the Laplacian Matrix (L) is seen in the following equation where it is RCxRC.

$$L = \forall(i,j) \in P \begin{Bmatrix} \deg(v_i), i = j \\ -1, i \neq j; v_i, v_j \mathit{ilebağlantılı} \\ 0, \mathit{diğeryerler} \end{Bmatrix}$$

Equality 1

(i,j) These are pixels in the P image. $\deg(v_i)$; is the degree of the i pixel and is the sum of all edges. Each line of L specifies how each i pixel depends on the others. In addition, the RC X 1 x vector that we define is the probability that the i pixel where line i is will be assigned as a specific label. $Q(V_i)$ represents the user-defined label of pixel I, for a specific user-defined k tag, the x vector is defined as follows.

$$\vec{x} = [x_1\ x_2\ x_3\ \ldots x_{RC}]^T\ ,\mathrm{s}_{.\mathrm{t}.}\ \mathrm{x}_i = 1\ \text{, eğer}$$

$$Q(\mathrm{v}_i) \neq \mathrm{k},\ \forall \mathrm{i} = 1,2,\ldots,\mathrm{RC}.$$

These pixels are not user-defined ($Q(V_i)$ except for the These are unknown and need to be resolved possibilities. If we re-edit the vector in Equation 2 into two sets, these pixels appear to be marked by the user first, Xm said, in the remaind there of this uncertainty, X_U , $x = [\vec{x}_M \vec{x}_U]^T$ equation. By following how we reorde the X vector, applying the same arrangements to L's lines, we get the following analysis of L.

$$L = \begin{pmatrix} L_M & B^T \\ B & L_U \end{pmatrix}$$

Finally, to calculate the unknown probability of the k. tag, it is important to resolve the following matrix equality.

$$L_U X_U = -B^T X_M$$

We create the x vector for each k tag and resolve the unknown possibilities of k. We'll find the maximum probability for x_U a pixel, no matter what label (mark). To use this to create a depth map, the method is changed as: In Random Steps, probability is in the range of [0 1]. Therefore, user-defined depths are allowed, and then depths are resolved for the rest of the image.X is allowed to manually adjust the probabilities of pixels marked with vector. The goal is just to make a depth solution for a label. For example, the user selects a value from [0 1] and intervenes on the image.0/1 means dark/light color or brightness. Then the rest of the depth is resolved by equality 3. The resulting possibilities can be used directly in depth to create stereoscopic 3D content. Scale-Space Random Walks (SSRW) has been studied in the method to increase sensitivity and reduce noise. This means that the image is sampled by a multi-resolution pyramid.

Random walkers are applied at any scale within the pyramid and merge using geometric mean. For edge weights, the dissimilarity function is a function often used in model classification:Sigmoidal function.The two pixels are defined as follows:

$$N(v_i, v_j) = \gamma \left(\frac{2}{1 + \exp\left(\beta D(\vec{c}_i, \vec{c}_j)^{\alpha}\right)} \right)$$

Equality 4

D(ci,cj)is the Euclidian distance of components between vi and vj. α,β.γ are parameters that control how different the two colors are. Experimentally, these parameters were α=γ=1 β=2

2.2.1.5 Graphic Separations for Images

Graphical separations are based on solving the Maximum-A-Posteriori Markov Random Area (MAP-MRF) marking problem with user input or harsh restrictions. The way to understand the solution to the problem is to understand how to find the most likely tagging for each pixel. This process is shown below. The solution can be found by minimizing energy function.

$$E(\mathrm{P}) = \sum_{p \in P} D_p(f_p) + \sum_{\{p,q\} \in N}^{n} V_{p,q}(f_p, f_q)$$

P here is the set of pixels that make up the image, and p is a pixel in the image. E(P) is the energy of the image. $D_p(f_p)$ data cost or a p pixel f_p represents the cost loss that occurs when it is assigned to the label. $V_{p,q}(f_p, f_q)$ is the cost of smoothness, or the cost loss (4-linked graphics in our system) when two different labels are assigned to two different pixels inside the spatial neighborhood of N. According to the chart theory, the maximum flow/minim separation chart is obtained, as shown in the minimized E(P) solution.

Efficient algorithms and software are obtained by minimization.

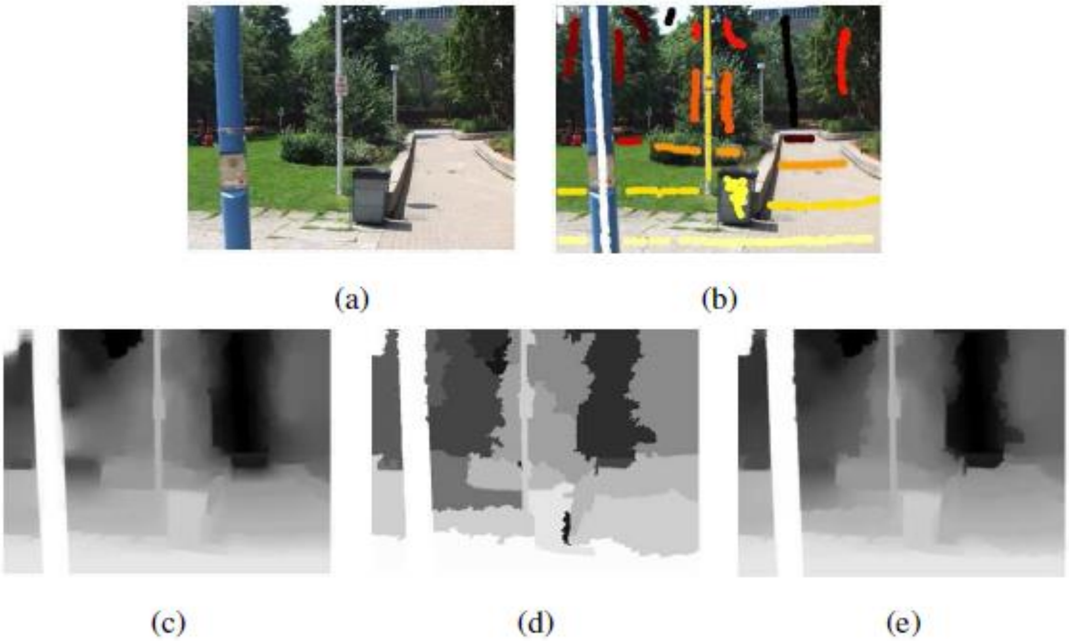
Depth map production can be considered a multi-label classification problem. However, chart separation focuses only on the binary classification problem. Similarly, each pixel has pre- and post-process costs associated with it. Therefore, each unique, user-defined depth value B 2 [1; N_d] is assigned as an integer. (N_d represents the sum of unique depths in user-defined marking)

Binary segmentation is applied individually for each label($b \in B$). User-defined marks with a b mark are assigned in the foreground while other user-defined marks are presented in the background. The rest of the pixels are what we want to mark. Chart separations N_d works for the sum of their time, first for each label, and the maximum flow values are saved for each mark on the chart in B. If a pixel is assigned only one mark, then the mark is assigned. However, if one pixel is assigned to more than one label, the label will be to the maximum maximum flow corresponding to the minimum amount of energy required to classify pixels. In some cases, even this does not yield results for each pixel. However, zone filling methods can be used to fix this.

In our graphics separation system, we use sigmoidal function to express the cost of smoothness (Equation 4). For data costs, we use the improved Potts model rather than the negative log histogram. This is because of the nature of colors that represent images. The probability of a negative log histogram is calculated in the RGB color space, which is inherently dashed. However, we would like to use the CIE L*a*b* color space, which inherently creates a better perception of human beings. In addition, CIE L*a*b* has floating-point values for each component, so it's not easy to determine how much space is required when creating a histogram. Table 1 shows how data costs are calculated when the user-defined sign bEB is selected in the foreground. In the table, a specific graphical separations are expressed as the fore fore fore front mark m selected during minimization. The K value is defined as follows.

$$K = 1 + \max_{p \in P} \sum_{q:\{v_p, \mathrm{v}_q\}} N(v_p, v_q)$$

Figure 4(d) shows a depth map obtained using only graphical distinctions. The same label in figure 4(a) is used to be consistent. As with Random Steps, only certain parts of the image were marked and a reasonably consistent depth map was created. But it should be known that there are parts of the image that do not understand any signs. However, this can be corrected by using zone padding methods. In the end, although a depth map is created where the object boundaries are very well determined, there is no change in the inner depth of the objects. If this had been viewed stereoscopically, the objects would have appear as "3D cardboard cutting." Random Steps create smooth depth curves, and with the difficult segmentation of chart separations, random steps set the groundwork for depth changes to make objects look more realistic. Graphics segments, on the other hand, eliminate overflow effect and precise limits.



2.2.1.6 Combining The Two

By bringing together two depth maps, depth prior is created. The initial depth is presented as an estimate, and this rough draft provides a depth information. This information is fed directly from the random steps algorithm. Depth priority is essentially the depth map in chart separations, and random steps should continue to help maintain strong boundaries in the depth map. Before we put it together, we need to modify the depth priority. Chart separations are represented by integers in the depth map [1.Nd], while the depth map of random steps is in the continuous range [0.1]. Both depth maps correspond to each other. But to be compatible, one must be converted to another. Information about which continuous label corresponds to which integer label is saved. The graphical separations are then completed. This match is the "map back" in the continuous range. As a result, the depth priority to combine the two depth maps feeds directly into the Random steps algorithm by changing the edge weights. But the edge weight equation is changed. The distance function is applied to include information from the depth first. If d_i represents the depth obtained from the depth of the i pixel, the edge weight equation is modified and defined by changing the Euclid distance after it is processed according to the lookup table:

$$D(\overrightarrow{c_i,c_j}d_i,d_j,|a)=\sqrt{[D(\overrightarrow{c_i,c_j})]^2+\alpha(d_i-d_j)^2}$$

$\overrightarrow{c_i}$ $\overrightarrow{,c_j}$ and $D(\overrightarrow{c_i,c_j})$:section 2-B1. Set to 0.5 value is the scaling factor that specifies how much the depth priority should contribute to the output, and the positive fact is constant. d_P , specifies the p pixel in the depth priority. However, the dynamic range of depth information from the depth first is in the range [0.1]. The components of the CIE L*a*b* color gap are significantly larger, as shown in Equation 7. These conditions excel at depth-priority conditions. However, the entire image or entire film frame is first converted to the CIELAB color space. Each CIELAB channel is normalized individually. Thus, each component in the range [0.1] is used in Equation 7.

The combined results in Figure 4(e) An instance is set on a value of α =0.5. Compared to Figure 4(c) and Figure 4(d), Figure 4(e) contains the most attractive states. Random steps depth map only has a texture and slopes that are difficult to detect; depth priority has consistent and noticeable limits for objects in the scene. The trees and shrubs in the original picture are now much better distinguished from the objects in the background and neighboring objects. In addition, unclassified regions or spaces remained in depth first. These gaps are filled with the production of the final depth map. The final depth map is obtained using the color information in the image, taking advantage of the depth first.

2.3 Getting 3D Images

I-)Two or More Image-Based Algorithms

I-a) Depth with Binocular Inequality Method

I-b) Depth by Deviation from The Room

II-) Single Image Based Algorithms

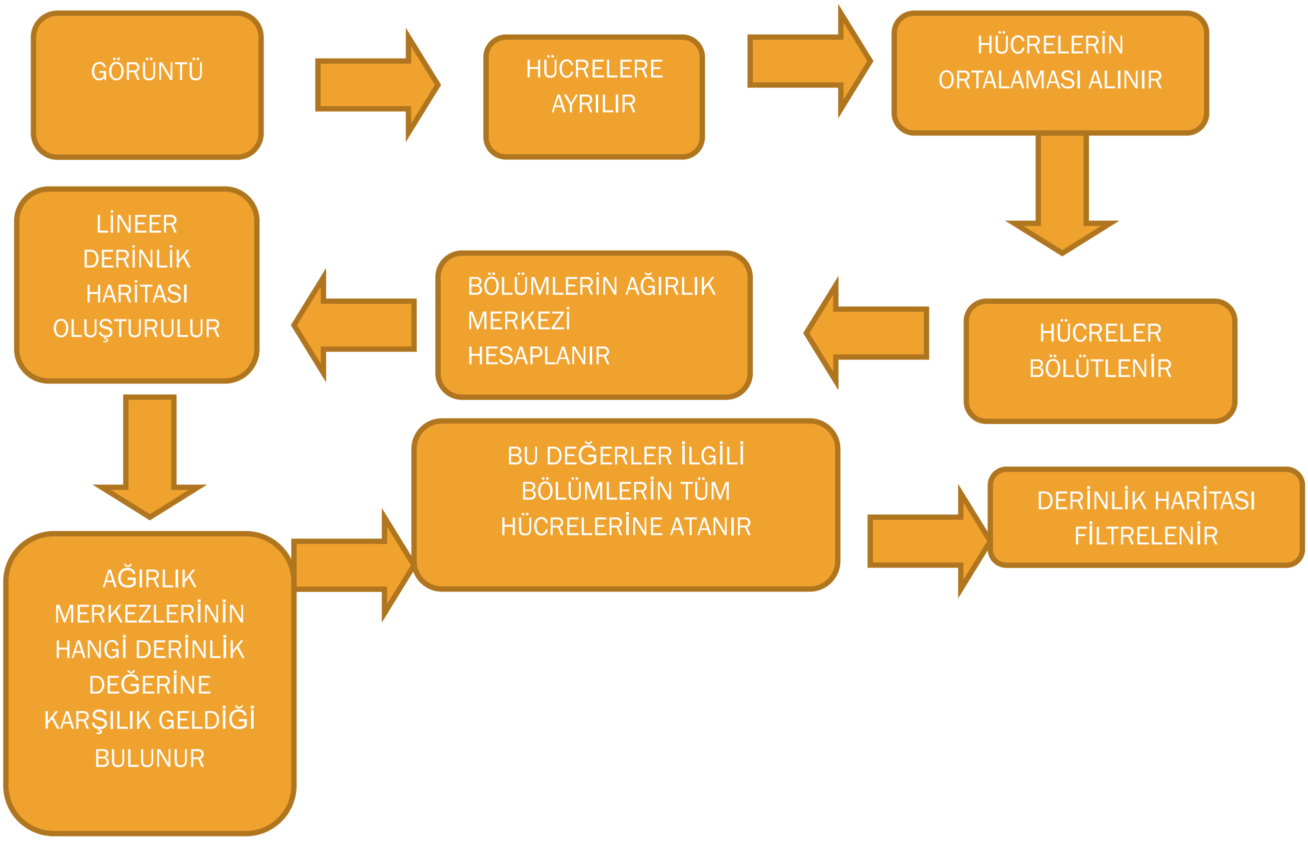
II-a) Depth with Blur Method

II-b) Depth with Focus Method

II-c) Depth from Geometric Perspective

II-d) Depth from Edge Knowledge

2.4 General structure of the applied method

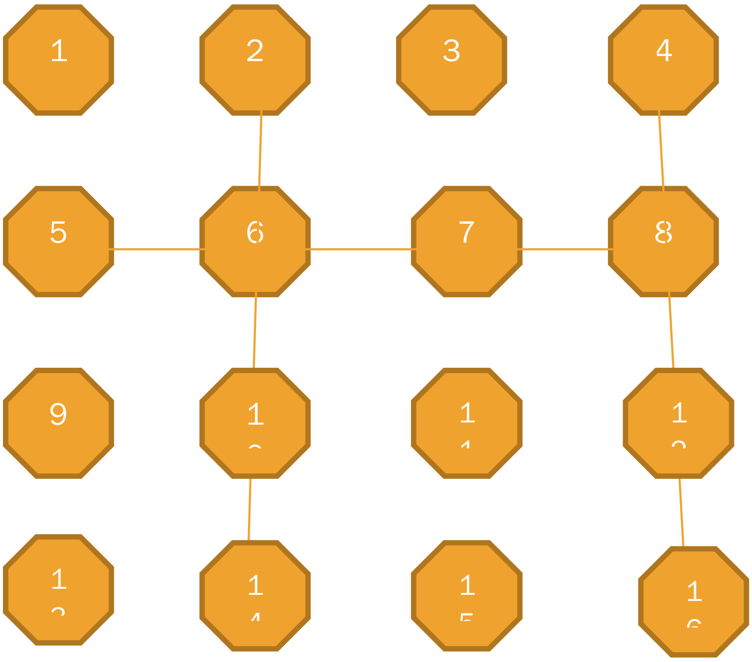

GÖRÜNTÜ
HÜCRELERE AYRILIR
HÜCRELERİN ORTALAMASI ALINIR
HÜCRELER BÖLÜTLENİR
BÖLÜMLERİN AĞIRLIK MERKEZİ HESAPLANIR
LİNEER DERİNLİK HARİTASI OLUŞTURULUR
AĞIRLIK MERKEZLERİNİN HANGİ DERİNLİK DEĞERİNE KARŞILIK GELDİĞİ BULUNUR
BU DEĞERLER İLGİLİ BÖLÜMLERİN TÜM HÜCRELERİNE ATANIR
DERİNLİK HARİTASI FİLTRELENİR

2.5 Split the image into cells

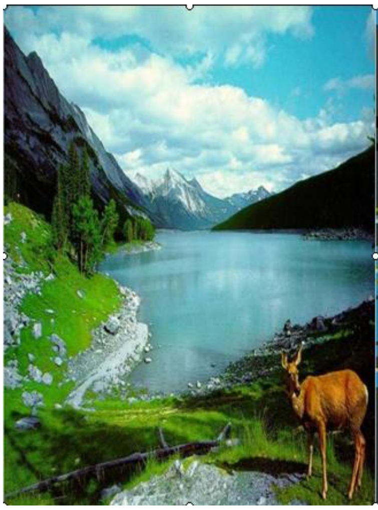
The purpose of separating the image into cells:

- Reduce computational complexity
- To get better color information

Original image divided into 4*4 blocks



Original image divided into 4*4 blocks



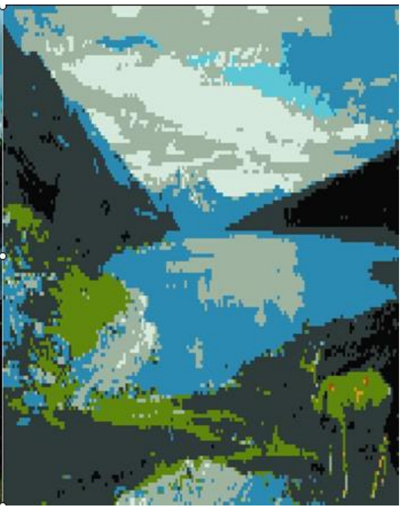
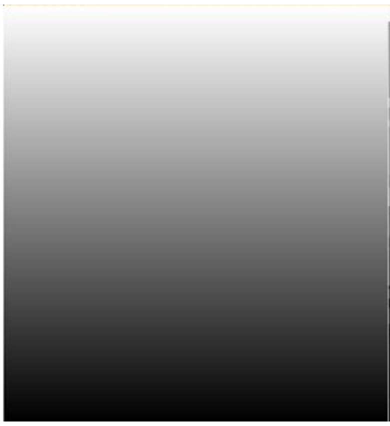


Figure 1- Original image Figure Figure 2-Divided image

2.6 Assigning a joint on a linear depth map



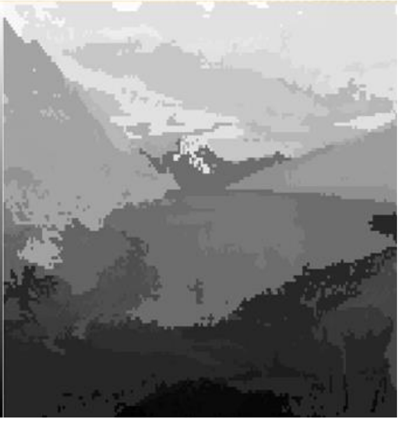
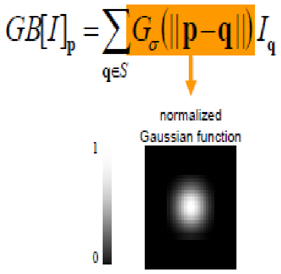


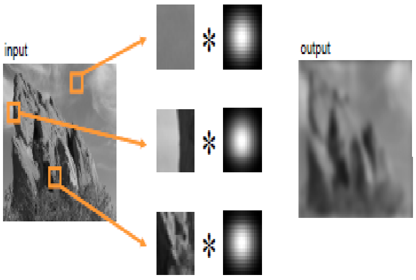
Figure 3-Linear Depth Map Figure Figure 4-Depth Map

2.7 Filtering the depth map

$$G_\sigma(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

*Gauss filter



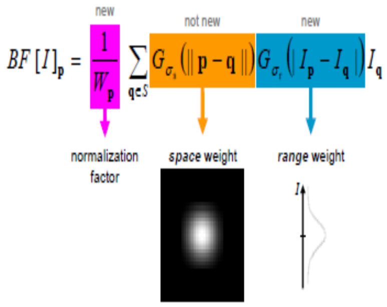


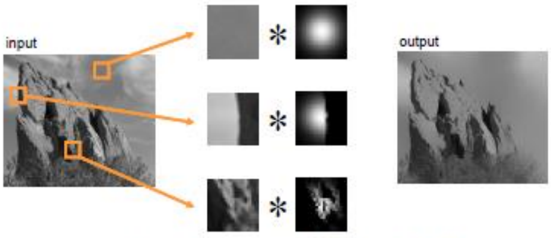
Same Gaussian kernel everywhere.

Filtering the Depth Map

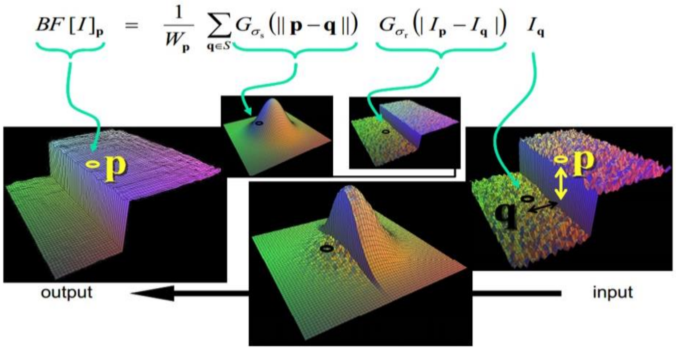
$$G_\sigma(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

□ Cross Binary Filter:

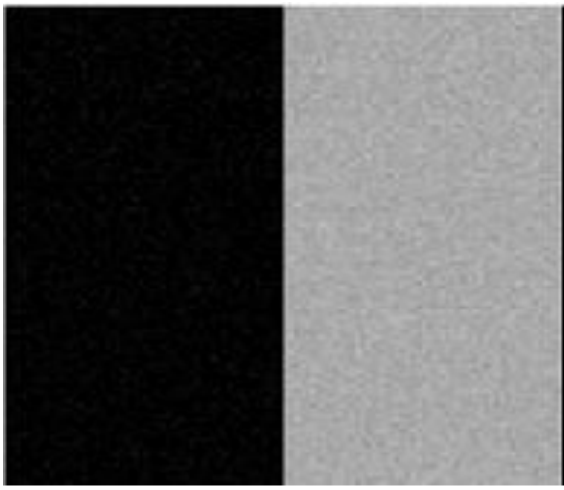

new
not new
new
$BF[I]_p=\frac{1}{W_p}\sum_{q\in S} G_{\sigma_s}(\|p-q\|)\,G_{\sigma_r}(|I_p-I_q|)\,I_q$
normalization
factor
space weight
range weight
I



The kernel shape depends on the image content.


$BF[I]_{\mathbf{p}} = \frac{1}{W_{\mathbf{p}}} \sum_{\mathbf{q} \in S} G_{\sigma_s}(\|\mathbf{p}-\mathbf{q}\|)\, G_{\sigma_r}(|I_{\mathbf{p}} - I_{\mathbf{q}}|)\, I_{\mathbf{q}}$
p
p
q
output
input

Filtering the Depth Map



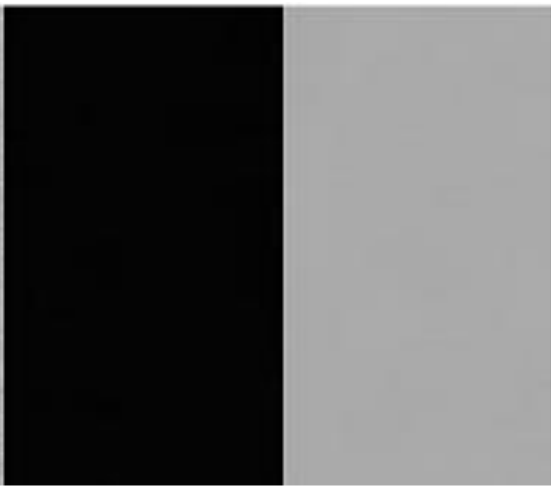
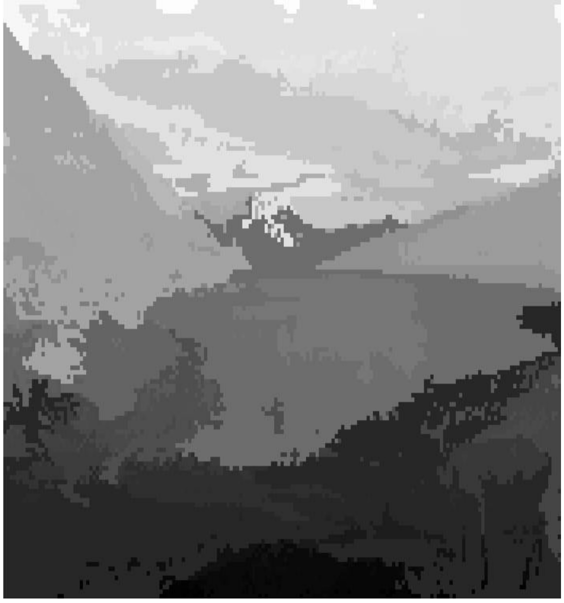


Figure 5-Unfiltered Image Figure

Figure 6-Filtered Picture with Diagonal Binary Filter



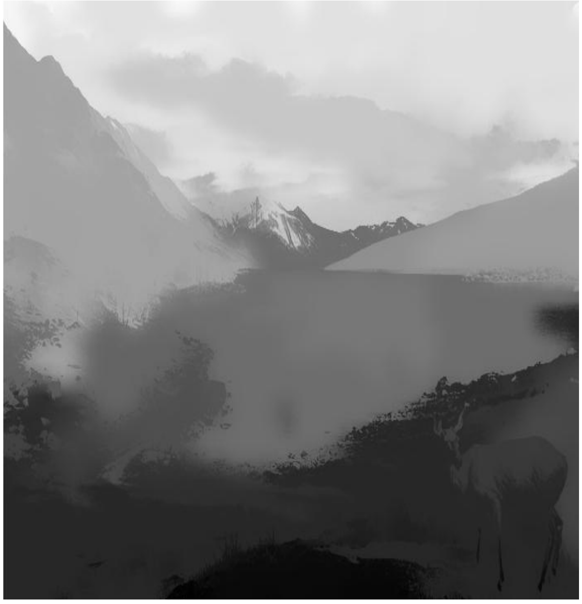


Figure 7-Filtered Image Figure

Figure 8-Filtered Picture with Diagonal Binary Filter

3. Results and Discussion

As a result of research and investigations, numerous methods have been found in the process of converting from two dimensions to three dimensions. It has been determined in the examinations that there are points where each method is weak and strong. The characteristic characteristics of the image were understood to be directly effective in selecting the conversion method to be used.

Conversion methods were examined and automatically and semi-automatically separated into two main heads. Today, many conversion methods are focused on automatic depth inerating, but reduced user intervention has made it difficult to control conversion results. Some errors have been seen to be unsolvable easily. It has been found that objects may not be able to get accurate feedback on depth information and some additional processes may be necessary. It has been found that user input used in semi-automated approaches to eliminate these disadvantages provides a significant advantage.

If the processes contained in the methods we recommend are applied correctly, quality stereoscopic images can be obtained and examined and seen in the experimental results of those who recommend the method. In the following period, critical information that will guide the software applications of these methods is absorbed and the methods are examined in theoretical detail.

References

[1] H. Wang, Y. Yang, L. Zhang, Y. Yang, and B. Liu, “2D-to-3D Conversion Based on Depth-from-Motion,” *Proc. IEEE Intl. Conf. on Mechatronic Science, Electric Engineering and Computer*, 2011.

[2] M. Guttmann, L. Wolf, and D. Cohen-Or, “Semi-automatic Stereo Extraction from Video Footage,” *Proc. IEEE Intl. Conf. on Computer Vision (ICCV)*, 2009.

[3] O. Wang, M. Lang, M. Frei, A. Hornung, A. Smolic, and M. Gross, “StereoBrush: Interactive 2D to 3D Conversion using Discontinuous Warps,” *Proc. Eighth Eurographics Symp. on Sketch-Based Interfaces and Modeling*, pp. 47–54, 2011.

[4] Y. Boykov, O. Veksler, and R. Zabih, “Fast Approximate Energy Minimization via Graph Cuts,” *IEEE TPAMI*, vol. 23, no. 11, pp. 1222–1239, 2002.

[5] C-. C. Cheng, C-.T. Li, and L-.G. Chen, “A 2D-to-3D Conversion System using Edge Information,” *Proc. IEEE Conf. On Consumer Electronics (ICCE)*, 2009.

[6] Yu-Lin Chang, Chih-Ying Fang, Li-Fu Ding, Shao-Yi Chen, and Liang-Gee Chen. Depth map generation for 2d-to-3d conversion by short-term motion assisted color segmentation. In Multimedia and Expo, 2007 IEEE International Conference on, pages 1958 {1961, July 2007.

[7] Tam, W.J. Liang Zhang, “3D-TV Content Generation: 2D-to-3D Conversion,” in *IEEE Intl. Conf. on Multimedia and Expo(ICME),* 2006

[8] C.-C. Cheng, C.-T. Li, P.-S. Huang, T.-K. Lin, Y.-M. Tsai, and L.-G. Chen, “A block-based 2D-to-3D conversion system with bilateral filter,” in *Proc. IEEE Int. Conf. Consumer Electronics*, 2009.

[9] Y.-L. Chang, et al, “Depth map generation for 2D-to-3D conversion by short-term motion assisted color segmentation,” in *Proc. ICME*, 2007.