3D-Dithered Ortho-Pictures: 3D Models from Independent 2D Images

Gershon Elber
Dept. of Computer Science, Technion – IIT, Haifa 32000, Israel
gershon@cs.technion.ac.il

Abstract
This work portrays a scheme to simultaneously 3D-dither two (or more) 2D gray-level images in \( \mathbb{R}^3 \), in orthogonal orthographic views, into one 3D model embedded in a cube, so that the different input images are seen from the different faces of the cube. From one axis-parallel orthographic view of the cube model, the first image is seen, and from a second, orthogonal, orthographic view, the second image is seen. We show that the dithering problem of more than one image does not have an exact solution as one image cannot be completely decoupled from the other; however, for images with rich enough gray-levels, the result will be a highly precise 3D-dithering of both images. Moreover, error correction methods common in classical dithering, such as the well known Floyd-Steinberg [8] algorithm, can be exploited in this 3D-dithering scheme. We then present some tangible examples, etched in glass (See Figure 1).

1 Introduction

The process of dithering a 2D image, as is described in any fundamental computer graphics book [9], follows two main steps, in general. In the first, a priori, stage, \( i/n^2, i = 0...n^2 \) gray-levels are encoded in 2D-dithering matrices of size \((n \times n)\). Each such dithering matrix encodes a gray-level between zero and one, of one pixel, and is created by painting the \( n^2 \) entries in the dithering matrix black or white. In this work, we henceforth denote the entries of the dither matrix as \textit{micro-pixels}. As an example, for \( n = 3 \), \( n^2 = 9 \) micro-pixels discretely create \( n^2 + 1 = 10 \) different gray-levels. In the ensuing discussion, we will interchangeably consider gray-levels to be between zero and \( n^2 \) or between zero and one, as \( i/n^2, i = 0...n^2 \), depending on the context, and whenever it is clear.

In the second stage of the traditional 2D-dithering process and given an image of size \((K \times K)\), colored or gray-level, each pixel is quantized to one of the \( n^2 + 1 \) discretely available intensity gray-levels. The closest quantized intensity is selected for every gray-level pixel and, possibly, the error between the original gray-level at the pixel and the quantized level can be computed. That said, this error can then be diffused or propagated to the neighboring pixels using, for example, the well known Floyd-Steinberg [8, 9] algorithm.

Given two images of size \((K \times K)\), this paper proposes a simultaneous 3D-dithering scheme that operates by meticulously positioning, typically \( K^2 \), voxels into a discrete cube volume of size \((K \times K \times K)\). Each voxel is further assigned a 3D-dithering matrix and then divided into \textit{micro-voxels}, again in a 3D micro-grid, of size \((n \times n \times n)\), where up to \( n^2 \) micro-voxels are set. Several video examples of different such etched glass cubes can be seen at http://www.cs.technion.ac.il/~gershon/V3dDithered.

---

1 Colors can be converted to gray-levels via the accepted CIE formula of \( \text{Gray} = \text{Red} \times 0.3 + \text{Green} \times 0.59 + \text{Blue} \times 0.11 \) [9]

---

Figure 1: 3D-dithering of two 2D images of the Obamas (left and right), etched in one 3D glass cube.
The end result of this arrangement is a 3D model, formed out of \( O(n^2K^2) \) micro-voxels, which looks like the first gray-level image from one view and like the second gray-level image from an orthogonal view. See, for example, Figure 1 that is an example of such a 3D dithered model that is etched in glass.

The rest of this work is organized as follows. Section 2 surveys the state of the art in this area. Section 3 presents the different stages of the basic 3D-dithering algorithm and in Section 4, some extensions are considered. In Section 5, additional results are presented and we conclude.

2 Previous Work

This work explores the synthesis of 3D-dithered geometry that projects to different gray-level images in different directions. The outcome has merit in both the sciences and the arts, with relevance to the areas of geometric modeling and the plastic arts, specifically. Hence, we now present related previous work that constitutes a nexus between the geometric modeling community and artistically-oriented design. This, while recalling that the vast majority of contemporary computer-based geometric modeling packages are geared toward mechanical design, focusing on the creation of 3D models that satisfy precise design and/or manufacturing needs.

Some artists specialize in creating models that look different from different views. Yaacov Agam [2] created 3D kinetic statues and used technologies such as lenticular printing. His Agamographs look completely different from different views, exploiting parallel vertical strips of different images that face different directions. Interestingly enough, the recent work of [3, 4] can be seen as an extension of Agam [2], creating shading that follows two input images from two different views by controlling the shapes of zero dimensional small elements like ellipsoids and boxes in [4] and pyramids in [3], instead of the one dimensional strips.

Shigeo Fukuda [10] created beautiful work such as “Duet”, “Love Story” and “Cat/Mouse”, all depicting two different scenes from two orthogonal directions. Other relevant artists are Markus Raetz [14] and Francis Tabary [16], who also sculpt such work; for example pieces showing one word from one direction, and another word, usually the antonym, from another. Also relevant is the “Escher for Real” project [6] that presents 3D geometry that looks completely different from different view direction; see Figure 2 for an example.

In [13], outline shadows are used as a source for synthesizing such 3D models, which projects to the different outlines in different views. Since the solution does not always exist, the problem is posed and solved as an optimization problem.

In [15], the problem of creating a 3D model that resembles two different shapes from two different views is also posed as an optimization problem of deforming a given model to follow the silhouettes of a different model. Figure 5 shows one result of [15]. Another noteworthy and somewhat related work is [12] where 3D Halftoning is defined toward tiling and filling of volumetric data sets such as porous materials.

The vast majority of this state-of-the-art work dealt with the silhouettes or outlines of the geometry and was incapable of handling gray-levels and shading. The noted exception is the work of Y. Agam [2] and his Agamographs and its extensions in [3, 4]. Indeed, our work herein advances this concept and explores the ability of presenting completely intermixed two or more gray-level images, as a one 3D dithered model.
We seek to derive an ability to simultaneously dither, as precisely as possible, two (or even three or more) 2D images of size $K^2$ pixels in a single 3D model—a cube of size $(K \times K \times K)$. The 3D cube model will be tiled by at least $K^2$ voxels, in 3-space, while typically, $K^2$ voxels are used. These $K^2$ voxels should cover all $K^2$ pixels of one image when projected along the $\vec{X}$ direction ($YZ$ plane) presenting the first image. In addition, when projected along the $\vec{Y}$ direction ($XZ$ plane) another set of $K^2$ pixels should be covered and the second image should be presented. Further, a third set of $K^2$ pixels may be covered, and if so a third image should be presented when the same set of $K^2$ voxels is projected along the $\vec{Z}$ direction ($XY$ plane).

The problem of placing $K^2$ voxels in a $K^3$ cubic volume so that they cover all $K^2$ pixels when projected along the $\vec{X}$ direction, along the $\vec{Y}$ direction, and possibly along the $\vec{Z}$ direction, has several simple solutions. For a coverage of the $K^2$ pixels in the $\vec{X}$ and $\vec{Y}$ projection directions only, one can simply place $K^2$ voxels along the diagonal plane of $X = Y$, at $(i, i, k), 0 \leq i, k \leq K - 1$. A more general random placement is also feasible via the following observation: Let $P = \text{Permute}(m)$ be some random permutation of the $m$ integers from 0 to $m - 1$. Then, Algorithm 3.1 will randomly place $K^2$ voxels in the $K^3$ cube so as to cover all $K^2$ pixels (i.e. constitute a coverage) when projected in the $\vec{X}$ or $\vec{Y}$ directions.

The fact that Algorithm 3.1 constitutes a coverage in the $\vec{X}$ or $\vec{Y}$ directions stems from the observation that in each row (z level) all indices of 0 to $K - 1$ are visited in both $X$ and $Y$. In Algorithm 3.1, we assume that every invocation of $\text{Permute}$ results in a new random permutation. Moreover, if the permutation is set to be the identity throughout, i.e. $P[i] = i$, we are, again, reduced to the diagonal plane, placing voxels in line 5 of the algorithm at $(x, x, z)$.

Covering all the $K^2$ pixels by $K^2$ voxels in 3-space, when projected in the three directions of $\vec{X}$, $\vec{Y}$ and $\vec{Z}$, is also feasible but is a bit more involved. One such feasible placement is along diagonal planes normal to vector $(1, 1, 1)$, as the set (of centers’ locations),

$$V = \{ (x, y, z) \mid (x + y + z) \mod K = 0, 
   x + y + z \neq 0, 0 \leq x, y, z < K \},$$

of $O(K^2)$ voxels which is also shown in Figure 3. Random placement of $K^2$ pixels that achieves the necessary coverage from the $\vec{X}$, $\vec{Y}$ and $\vec{Z}$ directions is also possible; see [7].

Having the $K^2$ voxels in place, each voxel is, in turn, formed out of $(n \times n \times n)$ micro-voxels ordered in a 3D-dithering matrix. When projected in the $\vec{X}$, $\vec{Y}$ (and possibly $\vec{Z}$) directions, this voxel will yield the desired gray-level of the relevant pixel in the relevant image, if possible. As a result, the 3D matrix will look like the two (three) given gray-level pixels of the input images, from two (three) different orthogonal views.

---

\[2\] From now on and unless otherwise stated, we will assume the placement of only $K^2$ voxels, which is the obvious minimum requirement while this number can optionally grow up to $K^3$ voxels, filling the entire 3D cube.

---

**Figure 3**: The ‘knishop’—a combination of a ‘knight’ (left) and a ‘bishop’ (right) as one merged smooth 3D model, following [13].
For now, we will assume a 3D-dithering process for two images only. Even this (relatively) simple problem of simultaneously 3D-dithering two images introduces new challenging degrees of freedom and is divided into several stages. In Section 3.1, we will show that for the matching process of pixels in the two images, for 3D dithering, not every gray-level of one pixel, from one view direction, can be a perfect match to every gray-level in a second pixel, in a second view direction. Equipped with the knowledge of what is feasible and what can be expected at the pixel–pixel matching level, in Section 3.2 we examine the matching a whole row of one image with a whole row of the second image. Then, the whole algorithm to 3D-dither two gray-level images in a single 3D cube is finally presented in Section 3.3.

### 3.1 Pixel–Pixel Matching

Consider two independent pixels of two different orthogonal images, already quantized to two (out of the \(n^2 + 1\) available) gray-levels, \(0 \leq n_x, n_y \leq n^2\). Having a 3D-dithering matrix of size \((n \times n \times n)\) of micro-voxels’ cells, we seek to place micro-voxels in the \(n^3\) cells so that when the matrix is projected in \(X\), exactly \(n_x\) micro-voxels will be seen (out of \(n^2\)) while a projection of the matrix in \(Y\) will reveal \(n_y\) micro-voxels (out of \(n^2\)).

Figure 5 shows one example for the case of \(n = 3\). A voxel with \(3^3 = 27\) micro-voxels’ cells is to be filled up with micro-voxels so that \(n_x\) micro-voxels are seen from \(X\) and \(n_y\) micro-voxels are seen from \(Y\). In effect, this creates a voxel with a gray-level of \(n_x/n^2\) from \(X\) and a gray-level of \(n_y/n^2\) from \(Y\). In a \(3 \times 3\) micro-grid, we can create ten \((n_x, n_y) \in \{0, ..., 9\}/9\) gray-levels. In Figure 5(a), three micro-voxels can create a coverage of \(n_x = 3\) and \(n_y = 1\). Figure 5(b) reveals a case of three micro-voxels’ coverage of \(n_x = 3\) and \(n_y = 2\), while Figure 5(c) presents a case of three micro-voxels’ coverage of \(n_x = 3\) and \(n_y = 3\). Possible placements of micro-voxels in these three arrangements are also shown as spheres in Figure 5.

Unfortunately, the difficulty is even greater and beyond this non-uniqueness. Not every gray-level of a pixel from \(X\) could be matched, using micro-voxels, with every gray-level of a pixel from \(Y\). Clearly, a coverage of \(n_x = 0\) could only be matched against a coverage of \(n_y = 0\). Moreover, a coverage of all nine micro-voxels (for \(n = 3\)), from \(Y\) (\(n_y = 9\)), requires a minimum of three micro-voxels covered from \(X\) (\(n_x = 3\)). See Figure 6.
A simple inspection of all possibilities of micro-pixel coverage in a matrix of size \( n = 3 \) (see Table 1) reveals what is possible and what is not. Impossible \((n_x, n_y)\) pairs are the result of close to complete micro-pixel covering in \( \vec{X} \) which cannot be concentrated in a few micro-pixel covering in \( \vec{Y} \), or vice versa. In other words, very bright images better not be matched with completely dark images. If an infeasible pair \((n_x, n_y)\) is forcefully matched, a close feasible pair \((n_x, n_y)\) must be selected, introducing some new gray-level Manhattan distance error,

\[
\epsilon = |n_x - \tilde{n}_x| + |n_y - \tilde{n}_y|.
\]

This \( \epsilon \) error can be minimized and/or diffused, a topic we will return to in Section 4.

The 3D-dithering matrices should be built only once. While the optimal (minimal) 3D-dither matrices are not unique, one can select a valid instance and expect a reasonable result. Motivation to alternate among the non unique optimal 3D-dithering matrices could stem from the desire to alleviate Moire patterns [1] in the result. In this case, several alternating dithering matrices could be stored for the same \((n_x, n_y)\) pair and used interchangeably.

 Nonetheless and while inherently exponential, a simple algorithm could be written to iterate over all \( 2^{2(n^3)} \) possibilities of 3D-dithering matrices of size \( n^3 \), only to select one (or more) instance of a 3D-dithering matrix for every feasible \((n_x, n_y)\) pair. In this work, we have precomputed and employed 3D-dithering matrices for \( n = 2, 3, 4 \).

Having an understanding of what can (and what cannot) be done at the pixel–pixel matching level, the next section looks at the matching of one complete row from one image with a complete row in a second image.

### 3.2 Row–Row Matching

At this point, it is clear that the simultaneous 3D-dithering of two images not only introduces errors due to the classic intensity quantization problem, but also due to the inter-dependency in the 3D-dithering matrices between the \( \vec{X} \) and the \( \vec{Y} \) directions, and the infeasibility of some \((n_x, n_y)\) quantized intensities of pairs of pixels.

Assume two images only, for now, that are inspected from the \( \vec{X} \) and \( \vec{Y} \) viewing directions. The pixels in a row of one image at some fixed \( Z \) level can only be matched with pixels in a row of the second image at the same \( Z \) level. Hence, we seek to form a one-to-one match between all the pixels in one row of the first image to all the pixels in one row of the second image; see Figure 7. Denote all the pixels in the row of the first image by \( I^1(p) \), \( p \in [0, ..., K - 1] \) and those of the second image by \( I^2(q) \), \( q \in [0, ..., K - 1] \). For each \( i \) in \( I^1(p) \), there exists a unique \( j \) in \( I^2(q) \) such that \( I^1(p) \) and \( I^2(q) \) are matched.

![Figure 7](image-url) | Two images are 3D-dithered in a 3D cube model, one row at a time (a). Here, the pixels of one (top level) row of one image, in red, are to be matched with the pixels of a second (top level) row of the second image, in green. Three matched pairs of pixels are presented in (b), as examples.
age by $I^2(q)$, $q \in [0, ..., K - 1]$. Several possibilities for such one-to-one matchings exist:

1. Enforce all voxels to be in one diagonal plane, which means the placements of voxels at $(p, p, z)$ $\forall p \in [0, ..., K - 1]$. That is, $I^1(p)$ is matched with $I^2(p)$. This is likely to generate the largest $\epsilon$ error.

2. Allow a completely free row matching, letting any one pixel in $I^1(p)$ uniquely match any one pixel in $I^2(p)$, and vice versa. This freedom reduces the problem to a weighted bipartite graph matching problem \cite{ref1} with the following constraints: pixel $I^1(p_1)$ ($I^2(q_1)$) must be matched to precisely one pixel $I^2(q_2)$ ($I^1(p_2)$) (only “monogamous marriages” are allowed). This is likely to generate the smallest $\epsilon$ error.

3. Allow a compromise between the above, 1 and 2, options. That is, $I^1(p)$ will be matched only with $I^2(q)$, such that $|p - q| < w$, where $w$ is some bound on the deviation, also denoted as the width of the matching. For small $w$, the 3D model is still bound to an almost diagonal plane. For $w = 1$, this case is reduced to option 1 and for $w = K$, we are reduced to option 2.

Much like \cite{ref2, ref3, ref4}, we assume a precise set of (orthographic) input views over the model. The larger the width $w$ is, the more distorted the presented image will be, when the 3D model is inspected from a finite distance or a somewhat different view. A partial remedy can be offered here to the influence of the distortion by limiting this $w$ width.

The solution to a weighted bipartite graph matching is well known in computer science \cite{ref5}. The Hungarian algorithm is used in this work (and hence no code is presented in this Section) to compute the optimal match, with a time complexity of $O(K^3)$ and actual real time performance of a few seconds for $K < 1000$. The weights here are set to zero for intensity matching pairs that have a valid 3D-dithering matrix (i.e. a ‘+’ in Table II) and the weights are set to be the Manhattan gray-level $\epsilon$ error, Equation (\ref{eq:distance}), if invalid (a ‘−’ in Table II). For example, and using Table II for 3D-dithering matrices of size $(3 \times 3 \times 3)$, the weight of pair $(4, 5)$, a ‘+’ in Table II will be zero but the weight of pair $(4, 0)$, a ‘−’, will be two as the distance error in Table II to the closest ‘+’ is two.

### 3.3 Image–Image Matching

We are now ready to combine it all together and create a 3D cube model of size $(K \times K \times K)$ formed out of $K^3$ voxels, with each voxel being formed out of up to $n^2$ micro-voxels. Algorithm 3.2 presents these top level steps. $\text{RowRowMatching}$ is the algorithm that matches two rows at level $z$, from Section 3.2, and returns the rows’ matched positions. $z$ is provided to $\text{RowRowMatching}$ so it can locate the matched points at the correct $Z$ level. Each voxel, with intensity pair $(n^m_x, n^m_y)$, is then tiled with micro-voxels, following the description in Section 3.4.

#### Algorithm 3.2 (3D image–image dithering)

**Input:**
- $K$: Output size to tile with $K^2$ 3D-dithering voxels;
- $n$: 3D Dithering size of micro-voxels;
- $I^1_{ij}$, $I^2_{ij}$: Two images of size $(K \times K)$;

**Output:**
- $C$: $K^2$ voxel positions, $P^m = (x^m, y^m, z^m)$, and intensity pairs, $(n^m_x, n^m_y)$, $m = 1, ..., K^2$, $0 \leq n^m_x, n^m_y < n$, to 3D-dither images $I^1_{ij}$ from $\vec{X}$ and $I^2_{ij}$ from $\vec{Y}$;

**Algorithm:**

1. $\mathcal{P} \leftarrow \emptyset$;
2. for $z = 0$ to $K - 1$ do
   3. $I^1_z \leftarrow I^1_{iz}$, $\forall i$;
   4. $I^2_z \leftarrow I^2_{iz}$, $\forall i$;
   5. $\mathcal{P} \leftarrow \mathcal{P} \cup \text{RowRowMatching}(I^1_z, I^2_z, z)$;
3. end for
4. $C \leftarrow \emptyset$;
5. for each $P^m$ in $\mathcal{P}$ do
   6. $C \leftarrow C \cup \{P^m, (n^m_x, n^m_y)\}$;
7. end for
8. Emit $C$,
4 Possible Extensions

The error between the desired (original) pixel intensity and the actual intensity can be accumulated for 3D-dithering of two images much like in the regular single 2D image dithering. Hence, classical error diffusion techniques can be used and are capable of alleviating the final global error, for example by adapting a variation of Floyd-Steinberg \[8, 9\] algorithm.

Because the width of the matching, \(w\), is typically greater than one pixel and since the expected pixel intensities in a row are set before the matching algorithm is applied, propagation of the dithering error to neighboring pixels is limited in this work to pixels in the next row only. The classic Floyd-Steinberg algorithm suggests the diffusion of the error that is introduced in pixel \((x, y)\) as in Figure 8(a) whereas herein, we purged the \((x + 1, y)\) direction, being in the same row, and instead use the table shown in Figure 8(b).

A central question throughout the discussion so far has been the extension of the presented solutions to three (or more) images viewed from the \(\vec{X}, \vec{Y}, \text{ and } \vec{Z}\) directions. Portions of the answer were already introduced when we showed how to tile the \(K^2\) volume using \(K^2\) voxels so that all the \(K^2\) pixels of the three images, which are viewed from \(\vec{X}, \vec{Y}, \text{ and } \vec{Z}\), are covered. With this ability, 3D-dithering matrices can also be designed with prescribed intensities from three orthogonal directions, a tedious task that again must be computed only once, delineating the feasible supported tri-gray-levels from those that are not.

The rest of the solution to the simultaneous dithering of three images is conceptually simple but computationally challenging. While the proper placement of \(K^2\) voxels in a cube so they cover three images from three different directions is feasible, a tri-partite graph matching, seeking the optimal matching (placement) for three sets, is to be applied. Unfortunately, the optimal solution to weighted tripartite matching is known to have exponential complexity. Only heuristic approaches could be expected to be feasible. See, for example, \[3\].

5 Additional Examples and Conclusions

We complete our presentation with two additional etched-in-glass examples, of pictures of a tiger and a lamb, in Figure 9, and of an Einstein portrait and his famous matter-energy conversion equation, duplicated many times, in the other image, in Figure 10. These glass cubes (i.e. Figures 1, 9 and 10) are 60mm on the side and contain an order of a million micro-voxels each, etched in glass using an existing focused laser beam technology.
We presented algorithms to derive the necessary set of points to 3D dither two orthogonal 2D images. Clearly this work can be extended in a variety of direction, having more than two images dithered together. Other directions of extending this work will be to relax the strict one-to-one matching, and the support of colors, although contemporary glass etching technology barely supports two colors.

6 Acknowledgments

This work was supported in part by the ISRAEL SCIENCE FOUNDATION (grant No.278/13). I would like to thank Tomer Vromen for implementing the Hungarian algorithm. I would also like to thank Lev Dvorkin and Gregory Gisinsky for their help in making the real tangible glass cubes.

References