Tyler Niles
May 2010
ENGG8840

AUTOSTEREOGRAMS AND THE DECIPHERING THEREOF

Introduction

The intent and focus of this undertaking was to establish a practical method of creating single image stereograms (popularly known as "magic eye" photos) from a given image, and a method of discerning the original image given a stereogram. This was accomplished by implementing computer programs (written in the C language) to process the images.

A process which creates a stereogram from a given image generates an autostereogram. There are several popular methods of creating such images. The process selected for this project is known as SIRDS: Single Image Random Dot Stereograms.

Deciphering a stereogram is a much harder process, and requires the foreknowledge of any of several parameters, depending on the method used. The popularity of solving this problem stems from the field of Applied Computer Vision and is better known as the "shape from stereo" problem.

The resulting C code and images (originals, stereograms, deciphered object images, and filtered images) are attached and labeled by filename. Furthermore, all source code was developed using the GNU C compiler in the Cygwin environment on a Windows platform. Since the main file format for processing was RAW binary, the libtiff library (www.*libtiff.org*) was integrated and data type conversion applications were written to enable viewing on the Windows platform.

Autostereograms

The main principle stereograms operate on is stereopsis, the seeing of two images with each eye. With the proper separation of the two images, human eyes can be forced to converge and focus beyond the physical image plane and thus reconcile both images into a single, seemingly three-dimensional images.

Normally, human eyes will converge and focus on a given object presented before them, as seen in

Figure 1. This is the typical way human vision operates. Since the eyes are separated by a certain Inter-Pupilary Distance (IPD), each image—while having much of the same content—is separated by a certain angle. The brain then takes each of these images and reconciles them into a single image, interpreting the known differences (from the inherent biometrics creating the separation of the images) as depth. This is how humans perceive depth and shape.

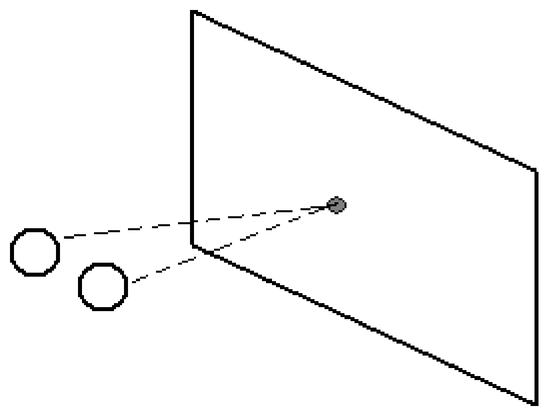

Fig. 1 – Eyes converging on image plane

A stereogram, therefore, takes a single image and duplicates it, and then calculates a separation for the original and the duplicate. Effectively, two copies of the original image are now stored within a single, new image (the stereogram). At this point, the image can then be overlayed with seemingly random noise. As long as the process is constrained in some fashion to preserve the two separated images, any random noise or repeating pattern can be applied to the image and provide a sort of camouflage for the separated images.

The separation is a variable calculated on the depth of the input object (see attached code). It is this separation of the two images which can then be perceived as depth. Since it is known that two images exist in the stereogram, any human viewing the stereogram can force their eyes to converge on the underlying hidden image (see Figure 2).

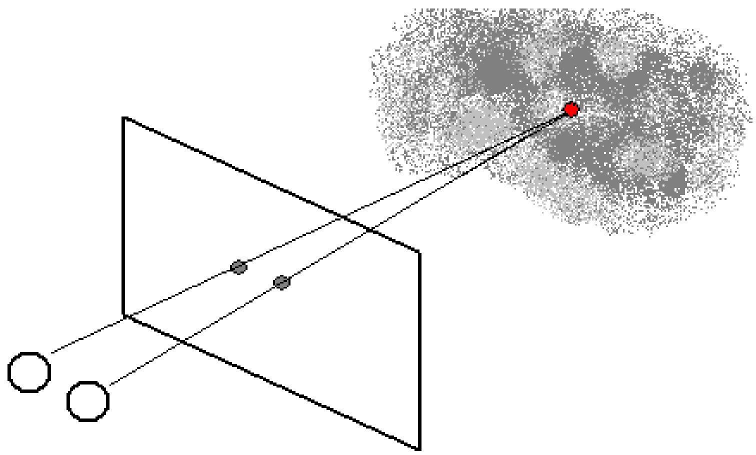

Fig. 2 – Eyes converging beyond image plane

The specific style of stereogram construction implemented in this project is a Single Image Random Dot Stereogram algorithm (*Thimbleby et al,* Displaying 3D Images: Algorithms for Single Image Random Dot Stereograms*, 1993*). Essentially, this calculates the required separation based on given geometric viewing constraints and records the new image with two copies of the original object embedded. During this process, random dots of black and white are written to every pixel of the image. If the pixel is constrained due to being part of the hidden image, the color between the left image and the right image must match. Otherwise, the pixel value can be randomly assigned.

Deciphering

As mentioned earlier, the real platform from which the popularity of and interest in these methods spring is the "shape from stereo" problem (*Kimmel,* 3D Shape Reconstruction from Autostereograms and Stereo, *Journal of Visual Communication and Image Representation, 2000).* That is, given two images—each from a separate camera, viewing the same object at two different angles, much like the human eyes operate—how best to determine the three-dimensional shape of the viewed object. This requires known parameters to describe the physical viewpoints from where the images come.

These parameters are best depicted by the inherent setup of human vision serving as an analogy. The cameras are like the eyes, receiving two separate images simultaneously. For a stereogram, perhaps the most important parameters are the near plane and the far plane of the image. That is, the minimum and maximum depths of the encoded (hidden) image.

Once these depths are known, the image can be scanned and rescanned at each unit depth interval while attempting to recognize repeating patterns (the two hidden, but matching, image elements) until all depth layers have been scanned. This recognition of the two image elements is also sometimes called correlation. Correlation functions exist which could provide improved performance in reconstructing the hidden image.

Results

The SIRDS algorithm employed proved to be quite robust. Images were successfully processed into SIRDS which can be used and enjoyed (see attached). These generated SIRDS were then used in turn as the test subjects of the deciphering application.

Originally, an algorithm called Minimal Area-based Correlation was intended for use in this project (*source*). However, it was found that the mathematics presented in this approach were insufficiently described. Instead, a simple pattern threshold method was utilized to recognize matching elements, assumed to be one of the two hidden images.

While this method led to much aggregate noise (improperly recognized patterns, not part of the original image, but coincidental strips), the majority of the original image was restored. Furthermore, additional filtering was applied to reduce some of this noise. This included 3x3 median filtering and an Adaptive Xu & Miller Rank 2-pass filter for comparison.

Conclusion

SIRDS methods are an efficient way to quickly implement simple stereograms. While several diverse methods of deciphering stereograms exist, it is clear that robust solutions present certain difficulties. It can be done, given the correct and known parameters—particularly the near and far planes representing minimum and maximum depth in the image. At this point, it is merely a matter of establishing a proper autocorrelation calculation and noise filter to generate a respectable interpretation of the original image.

Recommended further investigation includes a better (less noisy) pattern correlation process for reconciling the image elements, as well as developing an autocorrelation method for detection of the near and far planes in the image, so as to not require these as known parameters for the program.