2001

Image resizing and image quality

Michael Godlewski

SIMG-503

Senior Research

Image Resizing and Image Quality

Final Report

Michael Godlewski
Center for Imaging Science
Rochester Institute of Technology
May 2001

Table of Contents

Image Resizing and Image Quality

Michael Godlewski

Table of Contents

Abstract

Copyright

Acknowledgment

Introduction

Background

Methods

- Creating the Target Images
- The Psychophysical Experiment
- The Forced Choice Program: GFC

Results

- Statistical Results
- Questionnaire Results Combined with the Statistical Findings

Discussion

Conclusions

References

Title Page

Image Resizing and Image Quality

Michael Godlewski

Abstract

The need to resize an image is a common occurrence in all facets of industry, from data collection to consumer use. However, most of the research conducted in this area focuses on the various statistical measures, such as RMS deviation. While the statistical measures are suitable for determining the effectiveness of the interpolator f data collection, they do not necessarily coincide with the effectiveness of the interpolator in a consumer application. This study, therefore, focuses on the use of the three most common resizing algorithms, the bi- bi-linear and nearest neighbor interpolators, in the area of image quality from the consumer application perspective. The study is based on a psychophysical experiment designed to ascertain the preference of a particular algorithm for a particular type of image and determine what characteristics are used to form tl preferences. The results are presented so that they may provide direction in the use and development of interpolators for consumer applications.

Table of Contents

This report is accepted in partial fulfillment of the requirements of the course SIMG-503 Senior Research.
Title: Image Resizing and Image Quality
Author: Michael Godlewski
Project Advisor: Jonathan Arney, Ph.D.
SIMG 503 Instructor: Anthony Vodacek

Table of Contents

Image Resizing and Image Quality

Michael Godlewski

Acknowledgment

I would like to thank Dr. Arney for allowing me to pursue my research in the manner that I saw fit, and for believing in me when even I didn't know what I was doing. Secondly, I would like to thank that Center for giving me the opportunity to work with the great people that I met along the way. Finally, I would like to thank my parents, because with out them I wouldn't even be here at all, or I would have gone to a state school.

Table of Contents

Image Resizing and Image Quality

Michael Godlewski

Introduction

In nearly every field that uses images the topic of image resizing eventually must be addressed. Traditionally, this has been done by constructing mathematical metrics that measure the integrity of the data contained in the image after the resampling has occurred. By concentrating on the mathematical aspects of this type of image processing, the final products used in consumer imaging are somewhat forgotten. The mathematical measures usually used have no bearing on the final appearance of the image nor do they give any measure as to how a consumer would rate the acceptability of that image.

To attempt to close this gap, this research attempts to determine the interpolators that consumers would prefer given a certain type of image content. To do this a psychophysical experiment was devised to determine the suitability of the three most common interpolation algorithms that are used by typical consumers, the nearest neighbor, bi-linear and bi-cubic interpolators.

The objectives of this project are to provide a measure of the most suitable of the common resampling algorithms from a psychophysical standpoint. Using these conclusions it will become possible to determine which types of algorithms are appropriate for particular types of images based on their content.

Background

This approach to the measuring the types of resampling algorithms is rather unique because most of literature on the topic of resampling deals with metrics based primarily on mathematical calculations. Wh investigation does not replace the techniques of evaluating the usefulness of any particular algorithm, it does provide a measure that these mathematical techniques overlook. By not taking into account the preferences observers it is difficult to determine if the mathematical metric provides a good measure of true image quality from a standpoint of whether or not the image is pleasing to an observer.

By determining the correlation, if any, between observer preference and the resampling algorithm used resize an image, it will then become possible to use this information to focus the efforts of algorithm design to me the needs for particular image contents. Also, the information gained can be used when any image is required eventual use by an observer rather than as a source of data. It should provide guidelines for the type of algorithms to apply when resizing images with different types of content for use in a variety of applications.

Methods

Creating the Target Images

The first step was to create the image targets that would be used in this experiment. Adobe Photoshop was used to create the three targets, which were then saved in bitmap format. Each target was intended to simulate different aspects found in images that consumers would normally encounter. The first to be constructed was the "Grayscale" target.

The "Grayscale" target was intended to simulate fields of slowly varying value, or texture, in an image. The original target was created with a size of 100 x 100 pixels, and each patch, from left to right, was made to be darker then the previous by a pixel value of 16 in each of the three color channels, red, green and blue. The original was then resampled using the nearest neighbor interpolator to 200% of its original size, or 200 x 200 pixels, and saved in a bitmap format. The original was then resampled to 300 x 300 and 400 x 400 pixels and the resampled images were again saved in bitmap format. It is important to note that the original was always used as the basis from which the larger image was created to ensure that any artifacts that the interpolator introduced into the image were not compounded by resampling a previously interpolated image. This procedure was then used to create similar images for the bi-linear and bi-cubic interpolators, again always returning to the original and saving the results in bitmap format. The images created using these methods can be seen in the following figure:

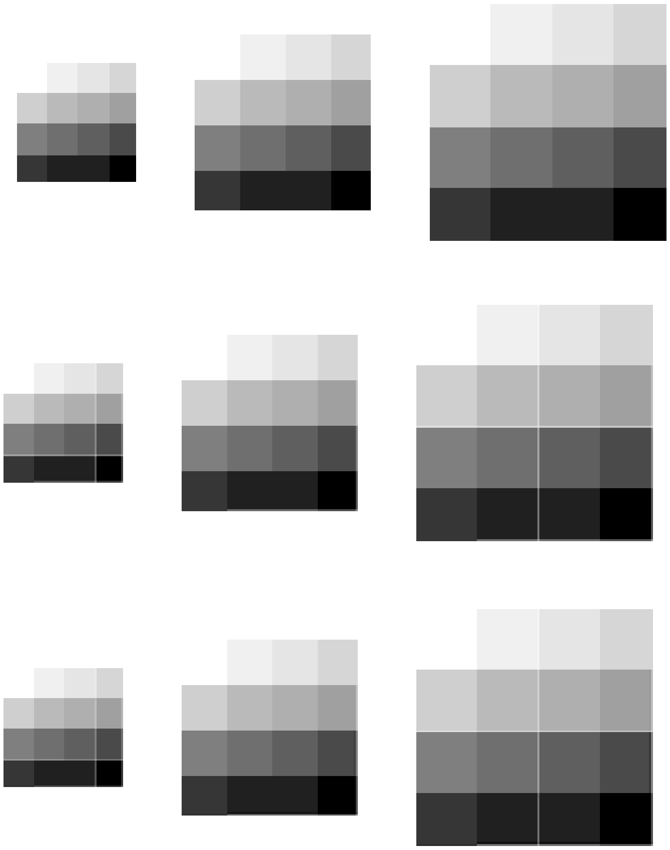


Figure 1: Grayscale Targets: From top to bottom, nearest neighbor interpolator,

bi-linear interpolator and bi-cubic interpolator

The "Star Target" image targets were intended to simulate the fine structures that can be present in image such as edges, blades of grass etc. The targets were created using the same methods that were used to cons "grayscale" targets, and can be seen in the following figure:

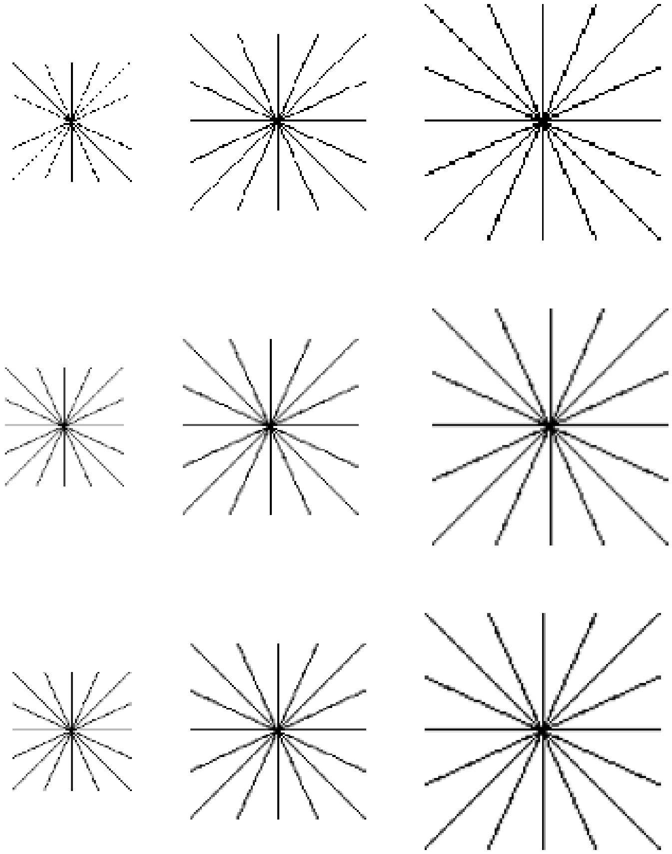
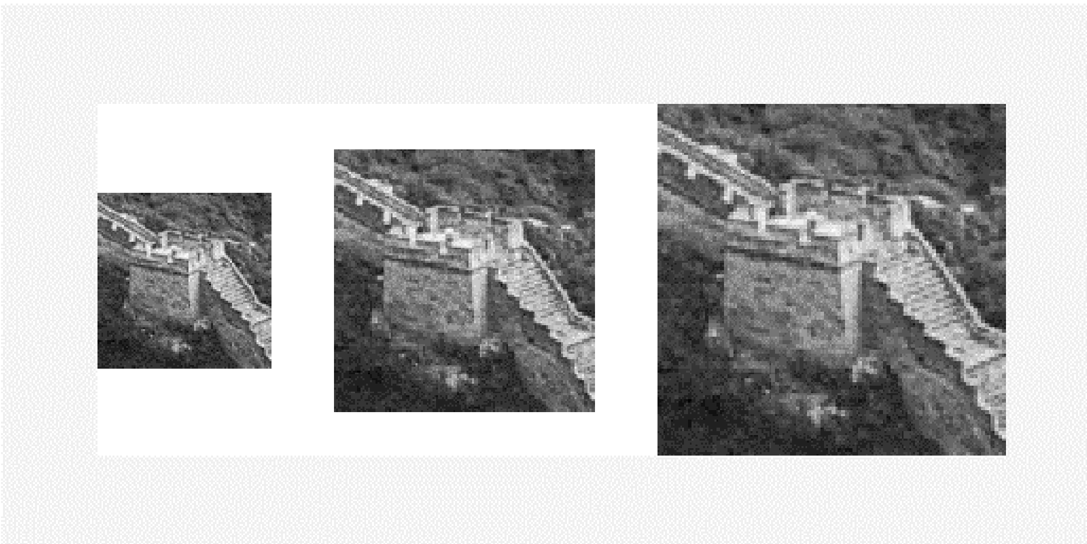


Figure 2: Star Targets: From top to bottom, nearest neighbor interpolator,

bi-linear interpolator and bi-cubic interpolator

The "Scene" targets were intended to simulate a stereotypical picture that a typical consumer might use targets were again constructed using the same methods as the two previous target categories, and can be seer following figure:



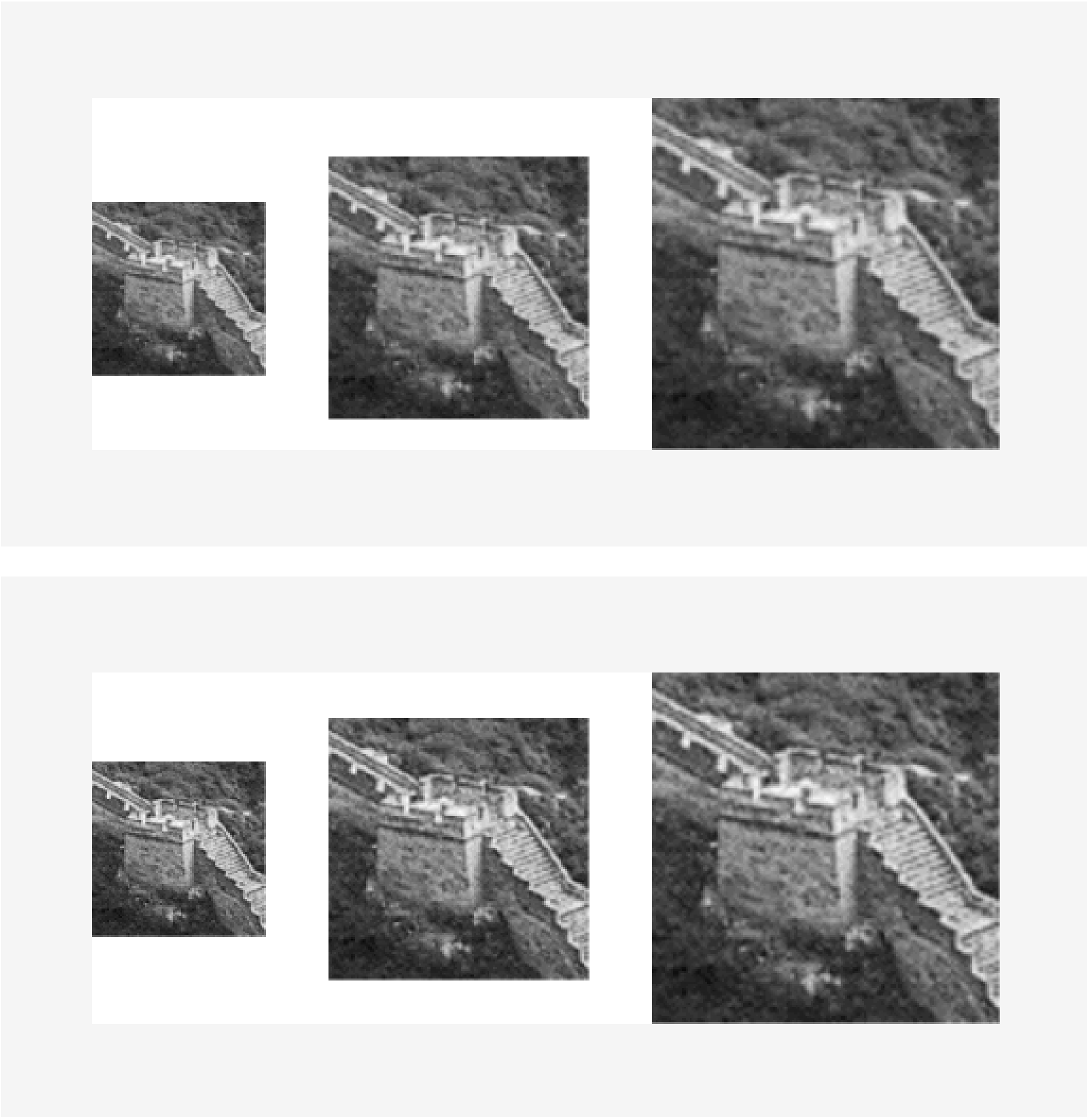


Figure 3: Scene Targets: From top to bottom, nearest neighbor interpolator,

bi-linear interpolator and bi-cubic interpolator

To see any of the images as they truly appear by category and group click on the desired group of images.

The Psychophysical Experiment

To determine the existence of a correlation, the psychophysical experiment was created using the guidelines as set up by Peter Engeldrum (1). The experiment was conducted along his guidelines and done in a manner that provided the highest possible amount of repeatability. The experiment was conducted in the same optics lab for each of the ten observers. The lab has approximately eighteen-percent gray walls and has overhead fluorescent lights. The computer monitor was angled so that no glare from the lights was present on the screen from the perspective of the observers, and the settings on it were never altered throughout the course of the experiment. The chair that the observers sat in was always the same distance from the monitor, so that the actual distance the observer was from the monitor only differed by the observer's own physiology. Observers were not allowed to change the position of the chair, but they were however allowed to lean closer to the monitor to allow a closer inspection of the targets.

The experiment itself was a two alternate forced choice type experiment, in which two target images were presented spatially, side by side, against an approximately eighteen percent gray background. Each target group was treated separately so that images of one group and size, the 200 x 200 grayscale targets for example, were compared against each other. That is the target created using the nearest neighbor interpolator was compared separately with the one created using the bi-linear interpolator and with the one created using the bi-cubic interpolator. Therefore each interpolated image was compared with the others in its group once yielding three comparisons per image size, per group, for a total of 27 comparisons per observer. The full experiment was conducted using ten observers and resulted in 270 total comparisons.

During the course of the experiment, after each image category was completed, the observers were asked to stop and answer the following question on a questionnaire provided:

Please list, by image group, the criteria that you used in determining your preferred image (*i.e.* line quality, sharpness, contrast, *etc.*).

This was done to determine the criteria that the observers used in choosing the particular images.

The Forced Choice Program: GFC

A program was written specifically written to perform the forced choice experiment for this project. It was written using Microsoft's Visual Basic programming language and was intended to be made available as freeware, and can be obtained here. The program allows for any bitmap image on the drive that the program is installed on to be used. The images chosen are displayed according to the order that they are initially placed in the image list. The 1st image is the one that will be displayed on the left of the screen and the 2nd in the list is the image that will be shown on the right of the screen, this pattern follows for the rest of the images in the list. Therefore every image that has an odd index in the list, 1, 3, 5 and so on, will be displayed on the left half of the screen and every even indexed image, 2, 4, 6 and so on, will be displayed on the right half of the screen. It has a feature that allows for some control over the positioning of how the images used are displayed on the screen, but this capacity is rather limited.

When the experiment is running, the observer uses the mouse to click on the image that meets the criteria set forth by the bounds of the particular experiment. After the image is clicked on the next two images in the list are automatically displayed and the process continues. This is done until all the images in the list have been displayed, and consequently creates the restriction that the number of images in the list to be displayed has to be even. Once all the images have been displayed the program returns to the image list setup screen. To retrieve the results there is a menu choice called "Results" which when chosen displays the all the images in the list using the full pathname and a zero or a one after it. The zero denotes that the image was not chosen and the one denotes that it was. The results then have to be copied by hand to some other format in order to perform any calculations that may be required.

Results

Statistical Results

The results obtained from the forced choice program were entered into Minitab 12 and placed into three groups according to category, "Grayscale", "Star Target" and "Scene". According to Devore (2), it was determined that the results had a multinomial distribution, and an appropriate test statistic was chosen, in this case the Chi squared statistic. A hypothesis test was devised to determine whether the patterns in the choices made are actually random. The hypothesis test is as follows:

p_1 = The probability of an observer choosing the Bi-Cubic interpolator

p_2 = The probability of an observer choosing the Nearest Neighbor interpolator

p_3 = The probability of an observer choosing the Bi-Linear interpolator

Null Hypothesis: H_o: p_1 = , p_2 = , p_3 =

Alternative Hypothesis: Ha: at least one p_i does not equal

Rejection Region: $\chi^2 \geq \chi^2_{0.05,2} = 5.992$

$\chi^2 = 48.6$ for the Scene Group

$\chi^2 = 32.27$ for the Star Target Group

$\chi^2 = 6.2$ for the Gray Scale Group

Since all the calculated χ^2
statistics were greater than the rejection value of 5.992, the null hypothesis must be rejected. Therefore the preferences of the observers were found to be not random. Because the results are not random, the data can be used to create estimates of the proportions of interpolator preference that a given population would have if undertaking this same experiment. The estimates of those proportions are as follows:

Scene Group	Star Target Group	Gray Scale Group
p_1 =	p_1 =	p_1 =
p_2 =	p_2 =	p_2 =
p_3 =	p_3 =	p_3 =

Questionnaire Results Combined with the Statistical Findings

The majority of participants responded that contrast and sharpness were the main determining factor in choosing the preferred image in the "Grayscale" group. Using the statistical findings, it is suggested that the nearest neighbor interpolator was most successful in preserving the desired characteristics. The other interpolators are effectively equal, which was expected since their appearance in general viewing conditions is indistinguishable.

For the "star target group", sharpness was the determining factor in choosing an interpolator. Again, the nearest neighbor interpolator was found to be the most effective, while the bi-cubic was found to be the second most effective. Finally, for the "scene" group the overall appearance of the image was used as the determining factor for the preferred interpolator. The interpolator that provided that most natural appearance was determined to be the bi-cubic interpolator. This result is not surprising as the bi-cubic interpolator is the closest approximation to the ideal interpolator.

Discussion

The questionnaire results when combined with the statistical results show that there is a measurable preference for interpolator type given image content. Some of these preferences are stronger than others, as is demonstrated by the proportions seen in the "Scene" and "Star Target" groups. The results from the "Grayscale" group require special attention. As can be seen by the low test statistic, which is rather close to the statistic for the rejection region, there would seem to be a weak preference for the nearest neighbor interpolator. This was due to the fact that under the viewing conditions of the experiment the images that used the bi-linear and bi-cubic were indistinguishable. They were not however identical, because under inspection on the order of a few pixels the images were in fact different. The white line artifacts that were in both sets of images had different variations under that closer inspection. It was determined, well after the experiment had been conducted, that the artifacts seen in those groups was due to the anti-aliasing feature in Photoshop being active during the creation of the test targets. This brings the results into question since when not using this feature the white line artifacts do not appear. When the images were created Photoshop 5.5 was used, and in this version of Photoshop it is not immediately apparent on how to disengage the anti-aliasing feature when performing an image interpolation. Therefore it is easily arguable that a normal consumer would not be aware of this problem and the artifacts would continue to occur. Although, in Photoshop 6.0 this option is either more easily accessed or is not a default feature and the artifacts do not occur. So the statistics generated for the "Grayscale" group in this experiment would only be applicable to images generated using a similar algorithm as the one used by Photoshop 5.5.

Conclusions

As is demonstrated in the population proportions and the evidence provided in the questionnaire, a typical consumer would have a preference for a certain image interpolation algorithm given a certain type of image content. The results suggest that for images, or portions of images, which have slowly varying fields of constant values that the nearest neighbor interpolator provides the best conservation of the desired characteristics. Those characteristics being sharpness and contrast. The same results were found for the "Star Target" group with sharpness being the most desirable characteristic. In the case of the "Scene" group, the bi-cubic interpolator was found to be the one most preferred with the bi-linear interpolator a close second. For the "Scene" group the interpolator chosen was the one that provided the most "natural" appearance, which was in fact the word used most often on the questionnaire for that particular group.

These results provide proof of what most likely has been known in many areas of industry for years, that the bi-cubic interpolator does not only provide the most common and practical ideal interpolator mathematically, it is also the most preferred common interpolator for use in a consumer context. The results can also be used to as a guideline for the creation of new algorithms used in for the consumer market, perhaps leading to the creation of adaptive algorithms that use already known and tested techniques.

All the goals that were set forth at the beginning of this project were met with satisfying results. However, more research into this area should be conducted to confirm the population estimates created and to gather more specific, detailed accounts of the "nesses" used as criteria by the observers for the preferred images. The experiment itself should also use a more diverse sample population, since this experiment used what can be considered trained observers. Also, the program created for experiments of this type is far from complete or perfect. It performs the tasks it was designed to do well, but the issues of ease of use, output of results and alignment of images all need to be improved. This of course will be difficult due to an unfortunate computer problem that erased the source code files. A usable executable does exist and can serve as a template for an improved version in the future.

Table of Contents

Image Resizing and Image Quality

Michael Godlewski

References

1. Engeldrum, Peter, G., Psychometric Scaling: A Toolkit for Imaging Systems Development. Imcotek Press, Winchester, Mass., 2000. (ISBN 0-9678706-0-7)
2. Devore, Jay, L., Probability and Statistics or Engineering and the Sciences: Fourth Edition. Brooks/Cole Publishing, Pacific Grove, CA., 1997. (ISBN 0-334-24264-2)