A Quantitative Comparison of Common Interpolation Methods - a study investigating common interpolation methods used in digital / multimedia forensic analysts

by Jim Hoerricks

Abstract

The interpretation of visual evidence relies upon the human vision/perception system. The creators of interpolation methods do not always factor how the results of interpolation will be perceived. This small scale study examines the interpolation of missing data in common CCTV formats from a perceptual quality standpoint in an attempt to answer the question, which is the better method of interpolation for digital/multimedia forensic analysts to employ.

Topic

If you ask the average digital/multimedia forensic analyst about the most appropriate interpolation method to use in forensic science casework, you'll likely hear a consensus response that Nearest Neighbor is the best, the most appropriate, the most defensible, and thus the most suitable. If you follow up with a simple question, why, you'll generally face a blank stare or an anecdote about testimony from a previous case.

In Strengthening Forensic Science in the United States: A Path Forward (2009), the US National Research Council (NRC) identified a fundamental problem in the development of "forensics" as a science. The simple matter that the NRC has sought to address is that there is very little science in many of the forensic sciences.

"Barry Fisher, Director of the Crime Laboratory of the Los Angeles County Sheriff's Department, has said, "We run the risk of our science being questioned in the courts because there is so little research." In 2001 Giannelli wrote, "In many areas [of forensic science] little systematic research has been conducted to validate the field's basic premises and techniques, and often there is no justification why such research would not be feasible." As Smith et al. note, the United States has a renowned higher education system, and many basic research discoveries relating to the forensic science disciplines have been made in academia. However, the forensic science disciplines suffer from an inadequate research base: few forensic scientists have the opportunity to conduct research, few academics are positioned to undertake such research, and, importantly, the funding for forensic research is insufficient. Others believe that the field suffers because the research initiatives being funded and pursued lack an overarching strategic plan."

At this point, it becomes important to define "forensic science." For this, I'll refer to *A Framework to Harmonize Forensic Science Practices and Digital/Multimedia Evidence. OSAC Task Group on Digital/Multimedia Science* (2018): "Forensic science is the systematic and coherent study of traces to address questions of authentication, identification, classification, reconstruction, and evaluation for a legal context." What is a trace? "A trace is any modification, subsequently observable, resulting from an event." When someone walks within the view of a CCTV system, they leave a trace of their presence within that system.

To Fisher's point (above), practitioners are faced with not knowing the answers to some very basic questions that may be presented to them in court. There are a lot of myths about how processes and procedures may work that are shared on-line and at conferences. But, little quantitative analysis exists to tests these myths. One myth, that the Nearest Neighbor inter-polation method is "the best," "the most forensically sound," or "the preferred option," is a consensus answer. These presumptions haven't been adequately tested for common CCTV evidence files. The community's answers presuppose a question, "how does one objectively form an opinion?"

The problem is exacerbated by the fact that software manufacturers aren't always forth-coming with the sources of their algorithms. For example, there is adequate anecdotal evidence to be found about how Adobe Photoshop's (PS) various proprietary interpolations perform in various circumstances (e.g.

enlargement vs reduction). But, the actual source documentation is not available for research. As such, this study will employ Amped SRL's FIVE (AF) software (build 13609) as the results generated in AF include reference(s) to the algorithms' source documentation. Photoshop, however, will also be used due to its popularity with the community.

This study will attempt to address the lack of knowledge about appropriate interpolation choices in a novel way. We will not approach the problem of "which interpolation method is best" from a one-size-fits-all standpoint. Neither will we address the problem from a purely mathematical or mechanical standpoint. We'll ignore the rhetorical as well; that Nearest Neighbor is the easiest to explain to the Trier of Fact. In our study, we will employ the methodology suggested by Wang, et al. (2004), to judge the resulting image's quality across four separate metrics, Sum of Absolute Difference (SAD), Peak Signal to Noise Ratio (PSNR), Mean Structural Similarity Index (MSSIM), and Correlation. MSSIM, as noted in Wang (2004), is well matched to visual perception and will be the primary metric used in the study. The other metrics may be better known to the community and are thus included as a reference. The presentation of demonstrative exhibits in trial asks the Trier of Fact to utilize their perceptual faculties in processing the presentation. A perceptual method of evaluation, the MSSIM, seems the most appropriate metric to employ in order to rank the interpolation methods tested.

Problem Statement

The perception exists that choosing "the wrong" interpolation method can negatively impact one's work in the digital/multimedia forensic sciences. The problem with this perception is that data is rarely cited in support of the analysts' opinions on the matter. Terms like "industry standard" and "consensus" are used, but these are meaningless. Thus, this small-scale study will test the question across four metrics in order to provide data that can inform discussions and decision-making processes, as well as further research.

Problem Background and Causes

The 2009 NRC report indicated a lack of an organized research agenda.

"A complete research agenda should include studies to establish the strengths and limitations of each procedure, sources of bias and variation, quantification of uncertainties created by these sources, measures of performance, procedural steps in the process of analyzing the forensic evidence, and methods for continual monitoring and improving the steps in that process."

Ten years later, an organized research agenda has yet to materialize. Thus, independent organizations like Praeceptory at Towcester Abbey are free to explore topics of interest and attempt to answer the most common questions in the digital/multimedia forensic sciences, establishing the strengths and limitations of procedures as well as quantifying variation and uncertainty, as this study seeks to do.

This study focuses on the problem of inter-polation and utilizes lower resolution video frames (2CIF / 4CIF / D1). The CCTV industry continues to release recording technology to market that utilizes these low-resolution settings to meet their marketing goals of 25/30 fps per camera and duration of storage targets. Higher resolutions tend to reduce frame rates as well as the amount of days of footage that can be stored on inexpensive storage media. Lower price points and continued industry support mean analysts will continue to see these formats in their evidence files.

Research Methods

The problem with questions of "best methods" is the fundamental question, "in relation to what?" In the case of an interpolation of a 2CIF CCTV file, the ground truth is difficult to establish, half the data is already missing. With this in mind, this study will utilize a convenience sample of two resolution types, full video frames with square pixels (e.g. 640 x 480) (N=4) and full video frames with non-square pixels (e.g. 704 x 480 & 720 x 480) (N=3). Given the small convenience sample of files tested, every effort was made to vary the content types within the files, e.g. day/night, colour/black and white, indoor/outdoor.

To begin, a copy of these files will have their resolution reduced by half via line deletion in Photoshop. In this process, every other line of resolution will be removed from the sample frames. In this way, we will have a ground truth file, a reduced resolution file, and files restored to the original resolution via the tested interpolation methods. The process will help us determine which of the interpolation methods get us the closest to our ground truth file's values – but from a perceptual standpoint. The resulting files will be judged across four separate quality-based metrics, Sum of Absolute Difference (SAD), Peak Signal to Noise Ratio (PSNR), Mean Structural Similarity Index (MSSIM), and Correlation.

The following interpolation methods will be tested:

- Nearest Neighbor (AF)
- Line Doubling (AF)
- Bicubic (AF)
- Area (AF)
- Lanczos (AF)
- Bilinear (PS)
- Bicubic Sharper (Enlargement) (PS)
- Preserve Details (Enlargement) (PS)

AF's implementation of Nearest Neighbor, Bicubic, Area, and Lanczos interpolation reference the following sources:

- Anil. K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall, pp. 253–255, 1989. ISBN: 0-13-336165-9.
- Anil. K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall, pp. 320–322, 1989. ISBN: 0-13-336165-9.
- Hsieh Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering", in IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 26, No. 6, pp. 508–517, December 1978. http://dx.doi.org/10.1109/TASSP.1978.116315

This study also uses Adobe's implementation of Bilinear interpolation, which is the same as AF's implementation of Bilinear (which also shares the above references).

AF's implementation of Line Doubling interpolation references the following source:

- E. B. Bellers and G. de Haan, "Deinterlacing - An overview", in Proceedings of the IEEE, Vol. 86, No. 9, pp. 1839–1857, Sep. 1998. http://dx.doi.org/10.1109/5.705528

Finally, this study uses Adobe's implementation of Bicubic Sharper (Enlargement) and Preserve Details (Enlargement). No references are available for these methods.

The question about appropriate interpolation methods comes up often in two basic scenarios:

- The restoration of missing data (e.g. 2CIF CCTV video files)
- Enlarging low resolution CCTV for display in trial (e.g. CIF to poster-sized enlargements)

This study will examine the former point, as we are able to construct an experimental design that includes the possibility for a known "ground truth image." The latter point will be the topic of a future study.

The tested single frames listed in Table 1 represent typical surveillance scenes. The tested single frames were all extracted from their respective data containers as bitmaps utilizing AF. Files 0 & 1 are sourced to a black and white camera-based system, depicting an outdoor street scene in daylight. Files 2 & 3 are sourced to an RGB colour system, depicting a dark indoor scene. Files 4 & 5 are sourced to an RGB colour system, depicting a well-lit indoor retail shop scene. Files 6 & 7 are sourced to a RGB colour system, depicting a well-lit outdoor street scene.

File No.	File Description	File Resolution
0	640 x 480 black and white	640 x 480
1	half 640 x 480 black and white	640 x 240
2	640 x 480 colour dark	640 x 480
3	half 640 x 480 colour dark	640 x 240
4	640 x 480 colour inside	640 x 480
5	half 640 x 480 colour inside	640 x 240
6	640 x 480 colour street	640 x 480
7	half 640 x 480 colour street	640 x 240

Table 1 - Images with Square Pixels

The similarity metrics for the files listed in Table 1 were computed in AF. The results are shown below in Table 2. Similarity Metrics in AF are accomplished by employing the Video Mixer filter. The Video Mixer filter allows the operator to overlay images from two processing chains. AF then computes and reports the similarity between corresponding frames using the metrics as shown below (Table 2). The methods with the best results are highlighted in yellow, with each image set's best results in bold text.

Comparison	**Interpolation Method**	**SAD (0 .. 255)**	**PSNR (dB)**	**MSSIM (0 .. 1)**	**Correlation (-1 .. 1)**
0 vs 1	Nearest Neighbor	0.832	35.2972	0.9796	0.994
0 vs 1	Line Doubling	**0.5836**	**44.459**	**0.9937**	**0.9994**
0 vs 1	Bicubic	0.8846	37.5008	0.9856	0.9964
0 vs 1	Area	0.832	35.2972	0.9796	0.994
0 vs 1	Lanczos	0.9059	37.3904	0.9839	0.9963
0 vs 1	PS Bilinear	0.8644	37.814	0.9862	0.9967
0 vs 1	PS Bicubic Sharper (Enlargement)	0.8672	37.8107	0.9861	0.9967
0 vs 1	PS Preserve Details (Enlargement)	0.8672	378107	0.9861	0.9967
2 vs 3	Nearest Neighbor	0.474	42.3548	0.9803	0.9923
2 vs 3	Line Doubling	**0.3889**	**42.8077**	0.9811	**0.9930**
2 vs 3	Bicubic	0.6431	42.6162	**0.9815**	0.9927
2 vs 3	Area	0.474	42.3548	0.9803	0.9923
2 vs 3	Lanczos	0.6609	42.5548	0.9809	0.9926
2 vs 3	PS Bilinear	0.6433	42.4955	0.9800	0.9925
2 vs 3	PS Bicubic Sharper (Enlargement)	0.6586	42.4567	0.9799	0.9924

2 vs 3	PS Preserve Details (Enlargement)	0.8212	41.5323	0.9744	0.991
4 vs 5	Nearest Neighbor	1.1429	33.5375	0.9745	0.9953
4 vs 5	Line Doubling	**0.7274**	**39.6587**	**0.9857**	**0.9989**
4 vs 5	Bicubic	1.4059	35.7759	0.9821	0.9972
4 vs 5	Area	1.1429	33.5375	0.9745	0.9953
4 vs 5	Lanczos	1.4027	35.8701	0.9821	0.9972
4 vs 5	PS Bilinear	1.3963	35.6361	0.9812	0.9971
4 vs 5	PS Bicubic Sharper (Enlargement)	1.4047	35.8397	0.9814	0.9972
4 vs 5	PS Preserve Details (Enlargement)	2.0894	32.528	0.9677	0.9942
6 vs 7	Nearest Neighbor	2.5245	27.1666	0.9131	0.987
6 vs 7	Line Doubling	**1.5843**	**31.0734**	0.9540	**0.9947**
6 vs 7	Bicubic	2.9107	28.7309	0.9312	0.9909
6 vs 7	Area	2.5245	27.1666	0.9131	0.987
6 vs 7	Lanczos	3.0039	28.6321	0.9294	0.9907
6 vs 7	PS Bilinear	2.9014	28.7117	0.9262	0.9909
6 vs 7	PS Bicubic Sharper (Enlargement)	2.9317	28.7982	0.9277	0.9911
6 vs 7	PS Preserve Details (Enlargement)	3.4135	27.9007	**0.9892**	0.9892

Table 2 - Comparison of Interpolation Methods for Images with Square Pixels

The tested single frames listed in Table 3 represent typical surveillance scenes. The tested single frames were all extracted from their respective data containers as bitmaps utilizing AF. Files 0 & 1 are sourced to an RGB colour system, depicting a dark indoor scene. Files 2 & 3 are sourced to an RGB colour system, depicting an outdoor scene with mixed lighting. Files 4 & 5 are sourced to an RGB colour system, depicting a well-lit outdoor street scene.

File Number	File Name	Resolution
0	704 x 480 inside dark	704 x 480
1	half 704 x 480 inside dark	704 x 240
2	704 x 480 outside mixed light	704 x 480
3	half 704 x 480 outside mixed light	704 x 240
4	720 x 480 outside daylight	720 x 480
5	half 720 x 480 outside daylight	720 x 240

Table 3 - Images with Non-Square Pixels

The similarity metrics for the files listed in Table 3 were computed in AF. The results are shown below in Table 4. Similarity Metrics in AF are accomplished by employing the Video Mixer filter as noted above. The methods with the best results are highlighted in yellow, with each image set's best results in bold text.

Comparison	Interpolation Method	SAD (0 .. 255)	PSNR (dB)	MSSIM (0 .. 1)	Correlation (-1 .. 1)
0 vs 1	Nearest Neighbor	0.2929	43.7352	0.9855	0.9973
0 vs 1	Line Doubling	0.2716	36.7522	0.9919	0.9865
0 vs 1	Bicubic	0.4249	39.5155	0.9884	0.9929
0 vs 1	Area	0.2929	43.7352	0.9855	0.9973
0 vs 1	Lanczos	0.4528	38.5906	0.9875	0.9911
0 vs 1	PS Bilinear	0.4057	38.9568	0.9887	0.992
0 vs 1	PS Bicubic Sharper (Enlargement)	0.4343	38.2291	0.9884	0.9904
0 vs 1	PS Preserve Details (Enlargement)	**0.2172**	**44.1484**	**0.9964**	**0.9976**
2 vs 3	Nearest Neighbor	1.2899	30.6976	0.9474	0.9893
2 vs 3	Line Doubling	**0.6601**	**36.5954**	**0.9819**	**0.9972**
2 vs 3	Bicubic	1.4705	32.5935	0.9633	0.993
2 vs 3	Area	1.2899	30.6976	0.9474	0.9893
2 vs 3	Lanczos	1.4942	32.5584	0.9627	0.993
2 vs 3	PS Bilinear	1.4268	32.7065	0.962	0.9932
2 vs 3	PS Bicubic Sharper (Enlargement)	1.4579	32.8221	0.9628	0.9934
2 vs 3	PS Preserve Details (Enlargement)	1.7181	31.9754	0.9602	0.9921
4 vs 5	Nearest Neighbor	1.3147	26.5165	0.9560	0.9835
4 vs 5	Line Doubling	**0.9511**	**27.8785**	**0.9711**	**0.9879**
4 vs 5	Bicubic	1.5075	26.9669	0.9628	0.9851
4 vs 5	Area	1.3147	26.5165	0.9560	0.9835
4 vs 5	Lanczos	1.5382	26.8908	0.9623	0.9849
4 vs 5	PS Bilinear	1.5011	27.1329	0.9619	0.9857
4 vs 5	PS Bicubic Sharper (Enlargement)	1.5345	26.9048	0.9610	0.9849
4 vs 5	PS Preserve Details (Enlargement)	1.9308	26.4668	0.9552	0.9833

Table 4 - Comparison of Interpolation Methods for Images with Non-Square Pixels

Conclusion

This study produced some rather interesting and unexpected results. From a perceptual quality standpoint, the Line Doubling interpolation method employed within AF is the clear leader. Line Doubling had the top MSSIM marks in the majority of comparative tests. For the two tests in which Line Doubling did not return the highest MSSIM value, it was a very close second. As for Nearest Neighbor, it regularly returned MSSIM results in the bottom third of the tested methods. Nearest Neighbor might be mathematically simpler to explain, as well as being easier to compute, but the results were farther away perceptually from the ground truth images.

Given the small convenience sample employed in this study, as well as the many software choices available in the world-wide marketplace, the results shouldn't be seen as representing a wider population. More research is needed, employing a larger sample not only of files but of software programs that may be utilized in the digital/multimedia forensic sciences. It's advisable that analysts not attempt to generalize based upon these results but to use this experiment's design to inform their own validations of their chosen interpolation methods.

References

1. Bellers, E.B., de Haan, G. (1998). Deinterlacing - An overview. Proceedings of the IEEE, Vol. 86, No. 9, pp. 1839–1857, Sep. 1998. http://dx.doi.org/10.1109/5.705528
2. Giannelli, P.C. (2001). Scientific evidence in civil and criminal cases. Arizona State Law Journal 103:112.
3. Hou, H., Andrews, H. (1978) Cubic splines for image interpolation and digital filtering. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 26, No. 6, pp. 508–517, December 1978. http://dx.doi.org/10.1109/TASSP.1978.116315
4. Jain, A. (1989) Fundamentals of Digital Image Processing. Prentice Hall, pp. ISBN: 0-13-336165-9.
5. National Research Council, Committee on Identifying the Needs of the Forensic Sciences Community. (2009). Strengthening Forensic Science in the United States: A Path Forward. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf

6. Neufeld, P. J. (2005). The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform. American Journal of Public Health, 95, 107–113. https://doi-org.wgu.idm.oclc.org/10.2105/AJPH.2004.056333

7. Organization of Scientific Area Committees on Forensic Science. (2018) OSAC Technical Series 0002R1. A Framework to Harmonize Forensic Science Practices and Digital/Multimedia Evidence. OSAC Task Group on Digital/Multimedia Science. http://dx.doi.org/10.29325/OSAC.TS.0002

8. Pyrek, K. (2007). Forensic Science Under Siege: The Challenges of Forensic Laboratories and the Medico-Legal Investigation System. Burlington, MA: Academic Press, p. 231.

9. Smith, F.P., Liu, R.H., Lindquist, C.A.. (1988). Research experience and future criminalists. Journal of Forensic Sciences 33(4):1074-1080.

10. Vincze, E. A. (2016). Challenges in digital forensics. Police Practice & Research, 17(2), 183. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=112454861&site=eds-live&scope=site

11. Wang, Z. Bovik, A.C., Sheikh, H.R., Simoncelli. E.P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600–612, Apr. 2004. http://dx.doi.org/10.1109/TIP.2003.819861

About the Author

Jim Hoerricks, PhD AVFA, Director of Research for the Praeceptory at Towcester Abbey, is a Certified Audio/Video Forensic Analyst (AVFA), is the author of the best-selling book Forensic Photoshop (available on Amazon.com), is a co-author of Best Practices for the Retrieval of Video Evidence from Digital CCTV Systems (DCCTV Guide), published by the Combating Terrorism Technical Support Office (CTTSO), is retired from police service where his unit (LAPD's Electronics Unit) was responsible for the collection and analysis of multimedia evidence (audio, video, images, and metadata from digital CCTV systems and mobile devices), and currently serves the Organization of Scientific Area Committees on Forensic Science (OSAC) as the Video / Image Technology and Analysis (VITAL), Video Task Group Chair; who moved the DCCTV Guide to an ASTM standard in 2018.