

Visual Comparability of 3D Regular Sampling and Reconstruction

Tai Meng, Alireza Entezari, *Member*, Benjamin Smith, Torsten Möller, *Member, IEEE Computer Society*, Daniel Weiskopf, *Member, IEEE Computer Society*, Arthur E. Kirkpatrick

Abstract—The Body Centered Cubic (BCC), and Face Centered Cubic (FCC) lattices have been analytically shown to be more efficient sampling lattices than the traditional Cartesian Cubic (CC) lattice, but there has been no estimate of their visual comparability. Two perceptual studies (each with $N=12$ participants) compared the visual quality of images rendered from BCC and FCC lattices to images rendered from the CC lattice. Images were generated from two signals, the commonly-used Marschner-Lobb synthetic function and a computed tomography scan of a fish tail. Observers found that BCC and FCC could produce images of comparable visual quality to CC, using 30% – 35% fewer samples. For the images used in our studies, the L_2 error metric shows high correlation with the judgement of human observers. Using the L_2 metric as a proxy, the results of the experiments appear to extend across a wide range of images and parameter choices.

Index Terms—visual comparability, perceptual quality, 3D regular sampling and reconstruction, Cartesian Cubic (CC) lattice, Body Centered Cubic (BCC) lattice, Face Centered Cubic (FCC) lattice

I. INTRODUCTION

A. Motivation

Sampled volumetric signals are widely used in such fields as biomedical imaging (computed tomography, ultrasound, magnetic resonance imaging, and microscopy) and computational sciences (fluid flow, astronomy, and biology). The common choice of sampling lattice (grid) is almost exclusively the Cartesian Cubic (CC) lattice since its structure is simple to understand. However, from the information and sampling-theoretic point of view the Body-Centred Cubic lattice (BCC) [19] and Face-Centred Cubic lattice (FCC) [11] have been proposed as superior approaches to sampling 3D signals. Each of these lattices captures as much of a signal as the CC lattice, while using about 70% as many samples as CC, and each has a theoretical advantage over the conventional CC lattice. The BCC lattice has been analytically shown [19] to be the optimal lattice for an isotropic, bandlimited signal that has been sampled without aliasing, whereas the FCC lattice has been analytically shown [12] to be the sampling lattice that introduces the least aliasing, when aliasing is unavoidable.

These analytic arguments represent an important case for BCC and FCC sampling over the conventional CC sampling, showing

Manuscript received XXXXX; revised XXXX.

T. Meng, B. Smith, T. Möller, and A.E. Kirkpatrick are with the Graphics, Usability and Visualization Lab at Simon Fraser University, Burnaby, BC V5A 1S6, Canada. E-mail: {tmeng,brsmith,ted,torsten}@cs.sfu.ca

A. Entezari is with CISE — Computer and Information Science and Engineering Department, University of Florida, Gainesville, USA. E-mail: entezari@cise.ufl.edu

D. Weiskopf is with VISUS — Visualization Research Center, Universität Stuttgart, Allmandring 19, 70569 Stuttgart, Germany. E-mail: weiskopf@visus.uni-stuttgart.de

that the same signal quality can be obtained with substantially fewer samples. These arguments are not definitive, however, for two reasons. First, the analytic arguments apply to signals that are more regular than actual signals. Real-world signals are rarely isotropic and bandlimited. Second, analytic arguments assess the impact of sampling on measures of fidelity to the original signal, whereas the ultimate measure of quality is the effectiveness of the final visualization after reconstruction and rendering. A medical visualization ought to provide the best support for diagnosis, while a visualization of experimental results ought to make patterns in the data as salient as possible for the scientist.

The gap between sampling and ultimate use is bridged by several steps. The sampled data is reconstructed into a three-dimensional scalar function, then rendered in a two-dimensional projection, and finally interpreted by the human observer according to the requirements of their task. Given these many steps, and their complexity, the signal-theoretic advantages of the BCC and FCC lattices may or may not translate into superior visualization effectiveness. But even if sampling is not the sole or final determinant of visualization quality, it is nonetheless crucial: The visualization cannot reveal any more data than remains after sampling. Later steps in the pipeline may introduce new distortions, but they cannot restore information discarded by sampling.

In this paper, we compare the effectiveness of visualizations derived from BCC and FCC sampling to visualizations derived from similar resolutions of CC sampling. We use a metric of end-point effectiveness, asking nonspecialists to compare images derived from the three sampling lattices. One of the signals chosen for sampling represents signals encountered in actual practice, while the other is a synthetic signal widely used to evaluate rendering. We test the hypothesis that the BCC and FCC lattices can produce visually comparable images using only about 70% as many samples as the CC lattice. This approach assesses whether the analytic estimates of sampling effectiveness [19] genuinely estimate quality perceptible to human observers.

We extend our perceptual results with numerical analyses. The constraints of experimental design limit the number of resolutions and source signals that we can ask our participants to compare. Participants can only review so many images before their judgment dulls. To cover a wider range, we first validate a numerical metric against our perceptual results, and then use the metric to extrapolate the experimental results to other signals and to a broader range of CC lattice resolutions. Another numerical comparison estimates the effect of data lost due to the original sampling of the source signals.

B. Contributions

Our main contribution is a pair of 12-participant studies visually comparing CC, BCC, and FCC sampling. For both synthetic

and computed tomography (CT) data, BCC/FCC sampling and reconstruction achieve visual comparability to CC sampling and reconstruction, using approximately 30% – 35% fewer samples.

This paper extends our previous study [16] of the visual comparability of BCC and CC with four additional contributions. First, we repeat our experimental protocol using the FCC lattice. Second, we improve the statistical analysis of our results for both studies by applying a nonparametric method. Third, for the same signals used in the perceptual studies, and the same reconstruction filters, we report the strong connection between L_2 errors and visual comparability. Finally, we use the results on L_2 error to estimate the impact of three potential sources of secondary effects: the downsampling pipeline, choice of sampling resolutions, and signal selection.

For simplicity of presentation, we will refer to both perceptual studies, the one on BCC described in our previous publication [16] and the one on FCC published here for the first time, as parts of a single study. Although the experimental sessions were separated by several months, the analysis reported here was done on both simultaneously.

II. RELATED WORK

Illustrated introductions to BCC, FCC, and CC lattices along with reconstruction filters are provided by Entezari et al. [5], [6], [7].

Mitchell and Netravali [17] performed the first user study validating reconstruction and rendering algorithms, estimating the perceptual effects of a class of C^1 and C^2 continuous, two-parameter, normalized cubic reconstruction filters on 2D samples. In their study, 9 expert users classified reconstructed images. Four reference images were shown, three exemplifying the properties of blurring, anisotropy, and ringing, while the fourth had satisfactory perceptual quality. Observers were shown images reconstructed using these filters, with randomly selected parameter values. For each generated image, observers were asked to select the closest reference image, implicitly classifying the perceptual quality of the underlying filter. The study demonstrated that numerically similar reconstruction filters can exhibit a wide range of perceptual effects.

Given the inadequacy of numeric metrics to predict perceived image quality, several perception-guided metrics have been proposed for 2D images. These metrics typically gauge the blurriness, sharpness, ringiness, or related characteristics of images. Marziliano et al. [15] define perceived blurriness in terms of average edge width. However, this metric cannot be applied to 2D renderings of volumetric signals because their perceptual quality cannot be measured by their amount of detail. For a volumetric signal, as the sampling resolution decreases, sharp features begin to disintegrate into tiny pieces upon reconstruction. This results in an increase in detail according to Marziliano et al.’s metric, when in fact perceived quality has degraded.

For a metric to predict the effectiveness of a visualization, it must distinguish good details from bad details. Even human experts find this distinction hard to articulate. Moreover, the use of different sampling lattices and different reconstruction filters, and the ensuing projection from 3D space to the 2D image plane make it hard to attribute “bad details” to any specific step in the process.

Perhaps more importantly, existing visual difference predictors [1], [9], [13], [15] measure the *visual similarity* to a reference

image, rather than the *visual comparability*. An image pair could be visually dissimilar to a reference image in different regions and yet be considered visually comparable by observers.

III. VISUAL COMPARABILITY EXPERIMENTS

We designed a pair of user studies to compare the three sampling lattices. The classic psychophysics protocols, which estimate a psychometric function [21], [22], are unlikely to apply because their underlying assumptions are not met. Unlike continuous measures of response intensity in psychophysics, our measure of user response, image preference, is discrete. Further, the factor of interest, perceived image quality, results from sub-factors, including curvature, symmetry, connectivity, and color; such multidimensional factors are unusual in psychophysics. Therefore, while our study employs the basic idea of a psychophysical study—observers’ subjective assessment of a stimulus—our task is adapted to our specific research question.

The studies each had 12 participants. The first experiment compared images generated from a CC sample with images generated from a BCC sample of the same signals, while the second compared CC and FCC images. Each study was a within-subjects design, with a single independent variable of BCC (FCC) sampling resolution. The experiments determined the relative sampling resolutions at which human observers found images rendered from BCC and FCC sampling visually comparable to images rendered from CC sampling. In each experiment, we created stimuli from two signals, one synthetic and one from an actual CT scan.

A. Choice of Task

We designed a task requiring the observer to choose between two passively-observed, two-dimensional images. We believe such a task is appropriate to an early study, and that its abstract nature makes it more generalizable than more specific tasks. Forced choices of “worse/better” are common in psychophysical studies. The task forces the observer to discriminate the two images carefully and choose by real or imagined differences. The degree of equivalence between two images is therefore indicated not by user report but by the degree to which observer preferences deviate from the chance level of 50%. If we had allowed observers to declare two images the “same”, we would have introduced observer-specific criteria of “equivalence”. Because such criteria are idiosyncratic and ultimately indeterminable (there is no prior metric of “perceptual degree of equivalence” to compare them with), it would have been difficult or impossible to aggregate results across observers. The approach we took, by contrast, is intended to allow combining results from many observers, and is widely used for that purpose in psychophysics.

Choosing a passive observation task eliminates potential confounds. An interactive task where the user manipulated a 3D view would make the results almost impossible to interpret because the design would introduce substantial nuisance variables. Different users would take different routes around the visualization, they would have different proficiencies with using the navigation tools (although readily-available tools support 3D navigation, the usability of 3D navigation remains an open problem), and the effort of working the navigation controls would introduce an attentional load, distracting participants from the primary task of comparing images. It would be impossible to usefully compare

two participants’ evaluations of two lattice samples because they might have looked at very different parts of the 3D space and taken very different lengths of time.

We also made a deliberate choice not to assess the impact of the different lattices on the performance of the kind of tasks for which such lattices are used. This study is a first step towards understanding the perceptual impact of the sampling lattices and correlating those perceptual results with error metrics. Although our simple task does not correspond, for example, to the way that radiologists might diagnose from a CT scan, our task in compensation allows us to draw on established approaches for analyzing and interpreting our results.

We emphasize again the simplicity of our design: All participants performed a forced-choice image matching task on the same sequence of images, without the distractions of running an interface. A simple design allows a simple interpretation. For a first study, simplicity of interpretation is crucial.

B. Detailed Task Description

For each signal, we generated several base images using the CC lattice. For each of these CC base images, we generated comparison images from several resolutions of BCC and FCC lattices. The software presented the observer with three images (Figure 1). A “reference” image, generated directly from the signal, was displayed at the top, while a pair consisting of a base CC image and a randomly-chosen comparison BCC (first study [16]) or FCC (second study, new for this paper) image was displayed at the bottom. Each image was 500×500 pixels in dimension.

The CC sampled image was randomly assigned to the bottom left or the bottom right slot, with the BCC or FCC image in the other slot. Participants were asked to click on the image from the bottom pair that “most closely resembles the image above”. If participants could not determine which bottom image more closely resembled the reference image, they were asked to make an arbitrary choice. Participants were informed of an “undo” button that would return the experiment software to the previous trial, in case they made a mistake.

Between trials, blank rectangular regions were displayed for 0.25 seconds, and flashed over the spaces where the new images were to appear. This eliminated an illusion of motion that had been observed during pilot trials.

Each study had two phases, a training phase and a main phase. The training phase gave the participant practice with the software and the image matching task. Image pairs for the training phase presented a fixed sequence of gradually more difficult choices. As with trials in the main phase, each pair contained a CC and a BCC/FCC image, and participants were not given feedback about the resolutions of the images they chose. Training images were taken from the set of images for the main phase. For the BCC vs. CC experiment, 9 images of ML and 4 images of Fish Tail were chosen for the training phase. For the FCC vs. CC experiment, 8 images of ML and 4 images of Fish Tail were chosen.

The main phase of each study consisted of 8 blocks of 24 trials each, for a total of 192 trials per participant. Trials alternated between blocks of ML images and blocks of Fish Tail images. A given study featured either all-BCC or all-FCC images for comparison with CC. Within each block, images were presented for all chosen sampling resolutions and camera views of that signal. Therefore, participants were shown each image pair 4

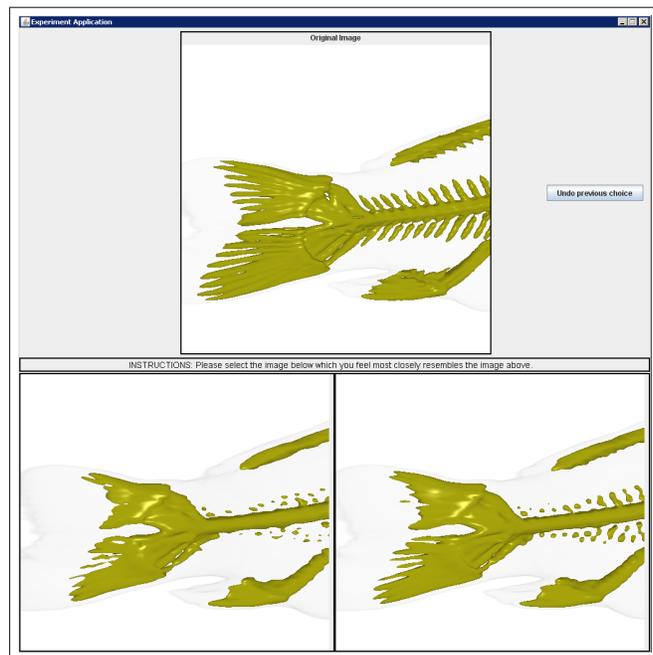


Fig. 1. Screenshot from the visual comparison task. The reference image is at the top, while the bottom shows a CC image and a BCC/FCC image. Participants clicked their mouse within the bottom image they found most comparable to the reference image.

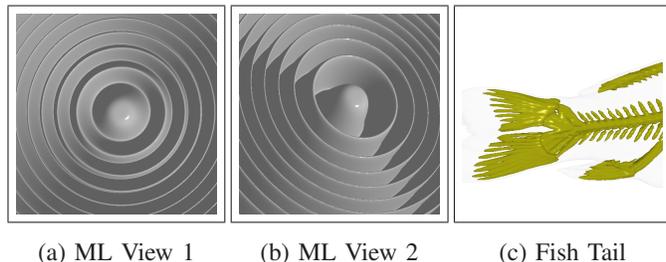


Fig. 2. Signals used in the visual comparability study: (a) synthetic ML signal (straight on view), (b) ML signal (tilted view), (c) Fish Tail CT signal.

times, with at least one block of trials between repetitions of the same image pair. Trials within blocks were randomized. To help alleviate boredom and maintain focus on the task, participants were encouraged to rest between blocks if they wished. Each participant took between 30 minutes and an hour to complete the user study.

C. Signal Selection

To avert participant fatigue, we restricted our stimuli to a list that could be comfortably reviewed in an hour. We chose two signals. The first is the Marschner–Lobb function (ML) [14] (Figure 2a–b), a common benchmark for 3D reconstruction algorithms.

The ML signal is a chirp-like pattern with a series of concentric rings. The further a ring is from the center of the ML signal, the higher its frequency content. Marschner and Lobb [14] originally proposed benchmarking 3D reconstruction filters using only the portion of ML within the domain of $[-1, 1]^3$, sampled at the resolution of $40 \times 40 \times 40$ on the CC lattice. This sampling resolution is sometimes referred to as the “critical sampling rate”.

The domain of $[-1, 1]^3$ captures a number of rings. However, it was not possible to show all these rings in an image without

also showing the boundaries created in the dataset by the range limit. Therefore, we sampled within the domain of $[-2, 2]^3$ so that the traditional domain of $[-1, 1]^3$ would have no boundary artifacts.

The second signal is a carp as recorded by a CT scan. This signal represents real world datasets from biomedical visualization. We focused on the tail (Figure 2c). The CT scan was presented as a CC lattice at a resolution of $256 \times 256 \times 256$. Because a CT scan is recorded in discrete form, we do not have an analytical description of the original signal.

To approximate the original signal, we reconstructed the CT scan using the tricubic filter described in Section III-H. We refer to the reconstructed signal as “Fish Tail”. This reconstructed signal is used both for rendering the reference images in the perceptual study and the numeric comparisons (Section IV-C).

Both ML and Fish Tail exhibit high frequency features along different orientations. Hence within one image, the perceptual effects of different sampling lattices on a wide range of frequency contents can be observed.

D. Downsampling the Fish Tail

Whereas it was possible to sample directly from the analytical definition of ML, it was necessary to reconstruct the Fish Tail dataset before sampling the reconstructed signal. We chose periodic interpolation followed by downsampling [18].

Using the $256 \times 256 \times 256$ CC volume data for Fish Tail as input, the downsampling process produces CC, BCC, and FCC data at reduced sampling rates. When downsampling the input dataset to CC datasets by a rational factor, we first upsample by zero padding in the frequency domain. Next, we downsample in the spatial domain by throwing away samples. Downsampling BCC and FCC began with the same upsampling step by zero padding the CC lattice, followed by downsampling from the CC lattice to a BCC or FCC lattice in the spatial domain [6].

E. Sampling Resolutions

To keep the number of trials feasible for a one-hour study, we used only three sampling resolutions for the reference signals. For the ML signal, we created a single $80 \times 80 \times 80$ CC reference sample over the domain $[-2, 2]^3$. This sample is denoted by CC80. For Fish Tail, we created CC reference samples using two resolutions: $140 \times 140 \times 140$ (denoted by CC140) and $180 \times 180 \times 180$ (denoted by CC180).

Figure 3 shows the BCC and FCC sampling resolutions as percentages of the fixed CC sampling resolutions. Note that the spacing of the BCC and FCC percentages in Figure 3 is inevitably irregular rather than equidistant, as the total number of points for each resolution is constrained to uniform sampling in 3D space. Such uniform samples can only be constructed for specific resolutions. The resolutions chosen for this study are the most closely-spaced resolutions that meet this uniformity constraint.

Pilot studies with expert observers (who did not participate in the actual studies) found that the BCC and FCC resolutions in the range 65% – 70% were most visually comparable to CC samples. Consequently, we included every BCC and FCC resolution within that range in the actual study.

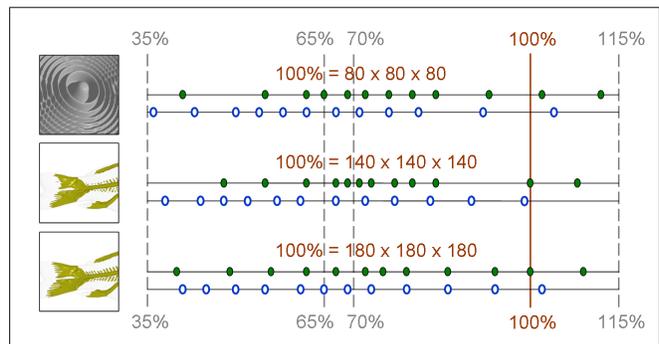


Fig. 3. Sampling resolutions for the perceptual study. Each reference CC resolution (in red) is fixed at 100%. The left column shows the CC reference images. BCC (solid green ovals) and FCC (blue rings) resolutions are represented as percentages of the reference CC resolution. The chosen BCC and FCC sampling resolutions are more densely distributed in the range 65% – 70%. The first row shows resolutions for the ML dataset at 80, and the second and third rows show resolutions for the Fish Tail dataset at 140 and 180 resolutions respectively.

F. Experiment Setup

The experiments were conducted in a small room that was generally insulated from outside distractions. The room was evenly lit with a fluorescent light fixture on the ceiling. The light fixture contained two 54 Watt fluorescent tubes with a white point of 3500K. Participants sat on a chair with an adjustable height of 45 cm – 55 cm, and were free to adjust the height of the chair. The chair was before a desk about 73 cm in height, and participants sat approximately 0.5 m away from a BenQ LCD monitor (model no. Q23W3). The monitor screen was set to a resolution of 1920×1200 pixels at a refresh rate of 60 Hz. The software ran under Windows XP on a desktop computer with an NVIDIA GeForce FX 5900 Ultra graphics card with 256 MB of memory, two Intel Xeon processors at 3.06 GHz each, and 4 GB of RAM.

G. Participant Selection and Administration

In total we recruited 24 participants: 12 of them compared BCC and CC images, and 12 of them compared FCC and CC images. No participant took part in both the BCC and the FCC experiments. All participants were graduate students from the School of Computing Science or the School of Engineering Science at Simon Fraser University. Age and gender were not considered to affect a person’s ability to detect relative differences in images, so no attempt was made to balance these variables across experiment conditions. Participants were paid CAD\$ 15.

We required that participants be unfamiliar with the process of sampling and reconstruction, and not involved in the generation of experiment images. Individuals with expertise in color science, medical imaging, and computer graphics in general were also excluded due to their potential expertise in perceptual research. Participants were required to have sufficiently good eyesight to perceive the stimuli clearly. Eyesight was evaluated by asking participants to read a short sentence displayed on the screen used in the study.

To minimize the possibility of influencing participants, the experiment administrators had minimal knowledge of the CC/BCC/FCC image generation process. At the start of the study, the administrators informed the participants of the nature of the experiment task, and then the participants signed a consent form.

During the training phase and the main phase, the administrators simply recorded participant comments.

H. Reconstruction Filters

The ideal reconstruction, in the space of bandlimited functions, involves a convolution with the sinc function. Since the sinc function has unbounded support, it needs to be windowed and truncated for practical (finite-time) purposes. Splines on the other hand, offer reconstruction filters that are local and have bounded support. Since any finite-length reconstruction filter has unbounded support in the frequency domain, practical (finite-time) filters exhibit some over-smoothing effects inside the passband and post-aliasing effects outside of the passband. While tensor-product B-splines can result in suitable reconstruction filters for the Cartesian lattice, the non-separable box splines have been demonstrated to be suitable for reconstruction on the BCC and FCC lattices.

The tricubic B-spline reconstruction filter was used for the CC lattice. The tricubic filter yields high quality images in part because it has C^2 continuity, which guarantees C^1 continuity of the normal vectors on isosurfaces. Such filters provide sufficiently smooth illumination and rendering [8], [20] for generating high-quality images.

For the BCC lattice, there is only one known C^2 -continuous reconstruction filter that provides high quality reconstruction: the quintic box spline filter. For the FCC lattice, there are no known comparable filter of C^2 continuity that yields high quality reconstruction. The 9-directional box spline filter, featuring C^3 continuity, was the next-best compromise. This choice introduces a mild confound to our comparison of the lattices because the C^3 -continuous filter for FCC reconstructs the signal more smoothly than the C^2 -continuous filters for CC and BCC. We suggest that this confound accords with actual usage, however, because we are pairing each lattice with the filter with which it is most likely to be used in practice. More details on the BCC and FCC filters can be found in [4], [7].

I. Camera View

The camera angle used to create an image might confound the effect of sampling lattice. Because the risk is greatest for ML, as it is a rotationally invariant data set, we generated ML images from two angles. First, a straight-on view, representative of biomedical imaging and presenting ML symmetrically, and second, a tilted view, more representative of visualization practice and presenting ML asymmetrically. By contrast, Fish Tail was only shown in a straight-on view, to keep the number of trials manageable within the one-hour limit of the user study.

J. Rendering Pipeline

All CC, BCC, and FCC sampled data were reconstructed using the filters described in Section III-H. After reconstruction, an opaque transfer function was applied to extract isosurfaces. The chosen isosurfaces provided an appropriate level of detail for the users to make comparisons. All ML images were rendered with one transfer function, while all Fish Tail images were rendered with a single transfer function, distinct from that for ML.

The central differencing estimator [3], [10] was applied to the reconstructed signal to estimate gradients. For each of ML and

Fish Tail, a fixed central difference step size of 1 unit, in the resolutions discussed in Section III-C, was used.

After the gradients were computed, Phong illumination [3] was applied with directional lighting. For each of ML and Fish Tail, the same illumination coefficients and lighting setup were used. Finally, the 3D renderings were projected onto the 2D image plane according to the camera views.

IV. PERCEPTUAL STUDY RESULTS AND DISCUSSION

A. BCC and FCC Image Preference

Figure 4 shows mean user preference data for BCC and FCC images, with bias-corrected and accelerated (BCa) bootstrapped confidence intervals [2] using 5000 bootstrap samples for each image. Given the comparatively small number of observers in each study, we chose a confidence level of 90%.

We define a BCC or FCC resolution to be visually comparable to a CC resolution when the confidence interval for the BCC or FCC resolution includes the 50% user preference line. We define the *range of visual comparability* as all resolutions falling between the two visually incomparable resolutions immediately before (after) the first (last) visually comparable resolution.

The results for ML View 1 and View 2 are shown in Figure 4(a–c). The results for Fish Tail CC140 and CC180 are shown in Figure 4(d–e). These plots show a general trend of increasing observer preference for both BCC and FCC images as their sampling resolution increases. The results also show that variability in participant preference increases as images become highly comparable. These are in excellent agreement with expectations.

The ranges of visual comparability (vertical dotted lines in Figure 4) are summarized in Figure 5. The ranges are generally 10% – 15% wide. Averaging the midpoints, we get 68% for BCC, and 65% for FCC images, within the experts’ range of 65% – 70% in the pilot study.

B. Qualitative

During the experiment, participants were encouraged to discuss the criteria they used to discriminate between images. Participants commented on a number of aspects of the discrimination task, none of which were unique to a specific lattice. The variable criteria used and comments made by participants highlight the complexity of the visual comparison task.

For both the ML and the Fish Tail signals, a few participants noted that the shading effects were different across different image pairings, but that they did not use this as the primary criterion for their discrimination. Some participants reported using a static set of criteria for quality, while others observed that their criteria changed during the experiment. Several participants noted that the task could be difficult in some cases: sometimes the two images were clearly different, but neither image was “better” than the other. The majority of participants also remarked that the discrimination task was easier for Fish Tail images.

For the ML images, participants remarked that they found curvature, symmetry, and the degree of distortion along edges to be important characteristics. For the Fish Tail images, participants mentioned focusing on the ribs of the fish and evaluating their connectivity and thickness compared to the reference image. Participants also commented that the shape of the larger, lower segment of the fish’s tail was a useful discriminator.

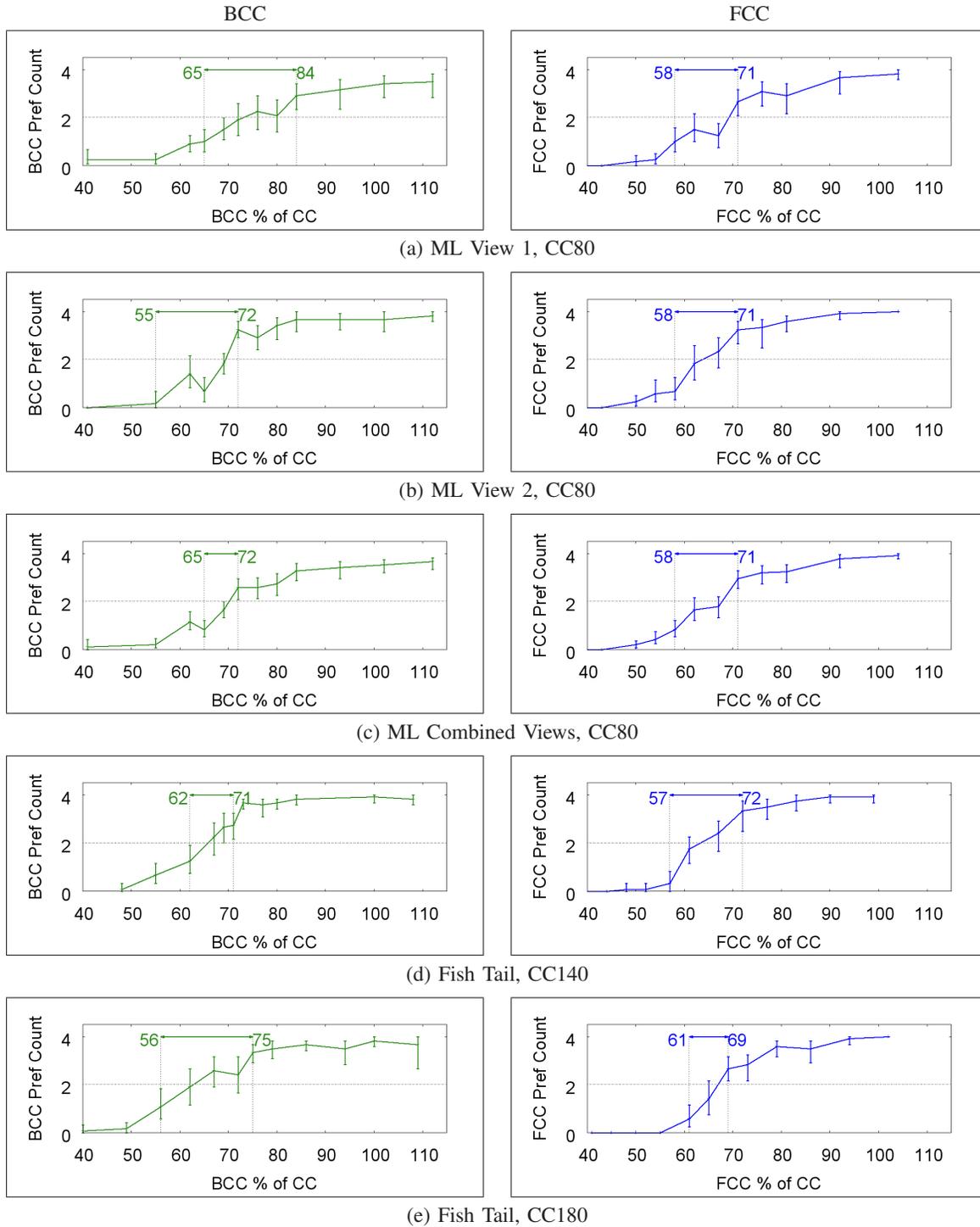


Fig. 4. Mean image preference ($N=12$ participants) per BCC/FCC resolution. Resolutions are constrained to the selected points by the BCC/FCC lattice. Error bars show 90% BCa confidence intervals. Vertical dotted lines show ranges of visual comparability (see text).

C. Numerical and Visual Comparability

As explained in Section II, existing visual difference predictors for 2D images [1], [9], [13], [15] do not address our research question. Fortunately, we discovered that a common 3D metric—the L_2 error metric—may be used to numerically estimate visual comparability. Although Mitchell and Netravali [17] argued that L_2 errors are in general not well related to visual outcomes from arbitrary reconstruction filters, they do relate well to visual comparability for our specific 3D reconstruction case.

Given functions $f, g \in L_2(\mathbb{R}^3)$, the L_2 error between them is defined as:

$$L_2(f, g) = \left(\int_{\mathbb{R}^3} (f(\mathbf{x}) - g(\mathbf{x}))^2 dV \right)^{1/2}.$$

We define the L_2 error of a sampled dataset as the L_2 error between the signal reconstructed from that dataset and its corresponding reference signal.

The L_2 error of a 3D dataset can be computed analytically because the reconstructed signals are piecewise polynomial functions. Deterministic quadrature rules are also possible options but, similar to the analytic computation, are computationally expensive specially in 3D. Therefore, we used a Monte-Carlo integration technique to obtain a tight estimate to the L_2 error. Since in our application the L_2 norm is merely used for relating the numerical errors to visual errors, we estimated the L_2 norm up to a precision level beyond which improvements were not observable.

We define *numerical comparability* when the L_2 error of a given CC dataset matches that of a BCC or an FCC dataset. Since each lattice has its own discrete (integer) resolutions (see Section III-E and Figure 3), the resolutions on each lattice do not exactly meet. A match to a CC dataset, at a given resolution, is defined by the two consecutive BCC/FCC resolutions whose errors are minimally below and above the error of the CC dataset (see for example Table I). We define *range of numerical comparability* to be the range of relative BCC/FCC sampling resolutions that are numerically comparable to some fixed CC resolution. To be consistent with the definition of visual range of comparability (Section IV-A), the range of numerical comparability is defined to be all relative sampling resolutions falling between the two numerically incomparable resolutions immediately before (after) the first (last) numerically comparable resolutions.

The numerical ranges of comparability are summarized in Figure 5 along with the corresponding visual comparability ranges. There is consistent overlap between the numerical ranges and their corresponding visual ranges, with the visual ranges tending to extend well below the numeric ranges. This confirms that L_2 numerical ranges are a useful approach to estimate ranges of visual comparability between CC and BCC/FCC sampled data under the conditions of our study, with numerical comparability tending to overestimate the number of samples required for visual comparability. Numerical ranges are employed in the next section to estimate the potential effects on our perceptual study of different parameter values.

V. EXTENDING THE PERCEPTUAL STUDY

In this section, we use numerical comparability to extend the perceptual study in two ways. First, we estimate the influence of two possible confounding effects, downsampling and choice of sampling resolutions. Second, we estimate the results of applying

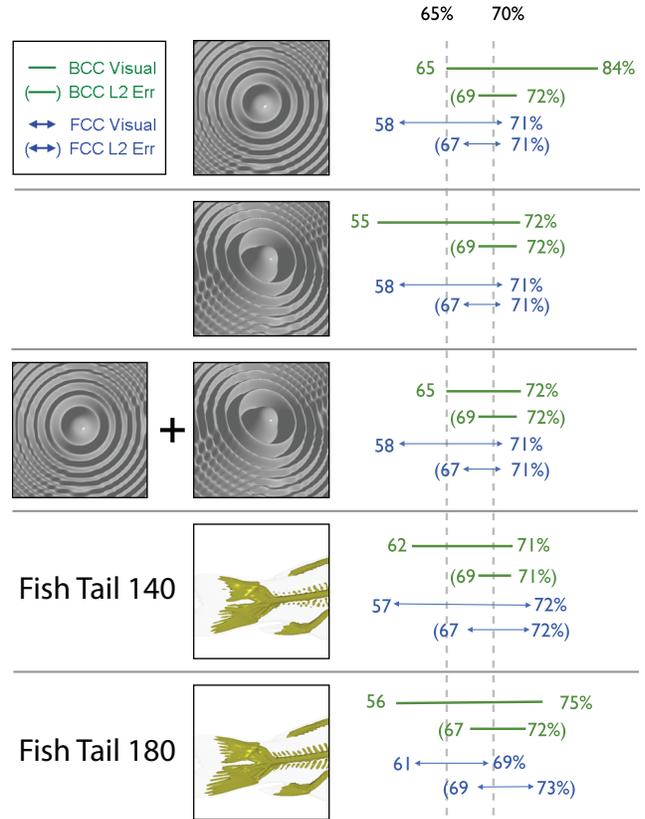


Fig. 5. Ranges of visual and numerical comparability. Left: Images from each condition. Right: Ranges for BCC (green, top) and FCC (blue, lower), with pilot estimates of visual comparability (gray dotted lines).

our perceptual study on different signals. In all cases, the effects appear small relative to the perceptual superiority of BCC and FCC sampling and reconstruction.

A. Checking for a Downsampling Confound

As discussed in Section III-D, the Fish Tail signal was downsampled because the original dataset was itself a CC sample of a continuous signal. In this section, we examine the numerical and perceptual effects of downsampling as opposed to direct sampling. We use the ML signal for the comparison because it is analytical and can be both directly sampled and downsampled.

The Fish Tail signal was downsampled from CC256 to CC140 (16% as many samples) and CC180 (34%). To generate comparable percentages for ML, the analytic ML signal was first sampled to CC128. The CC128 sample was next downsampled to CC70 (15% as many samples as CC128), CC80 (24%), and CC90 (34%). For each downsampled CC resolution, there was a corresponding range of downsampled BCC and FCC datasets.

The results of numerical comparability are summarized in Table I. The numerical ranges are nearly identical for reconstructions of direct and downsampled signals. This is also supported by simple visual inspection. Images rendered from downsampled data are as visually comparable to the analytic reference signal as images rendered from directly sampled data (Figure 6). Downsampling appears to have negligible perceptual effects.

TABLE I
EFFECT OF DOWNSAMPLING PIPELINE ON NUMERICAL RANGES OF
COMPARABILITY FOR ML DATASET.

ML	Direct Sampling	Downsampling
BCC vs. CC70 FCC vs. CC70	67% – 69%	69% – 71%
BCC vs. CC80 FCC vs. CC80	69% – 72%	69% – 72%
BCC vs. CC90 FCC vs. CC90	69% – 70%	69% – 70%

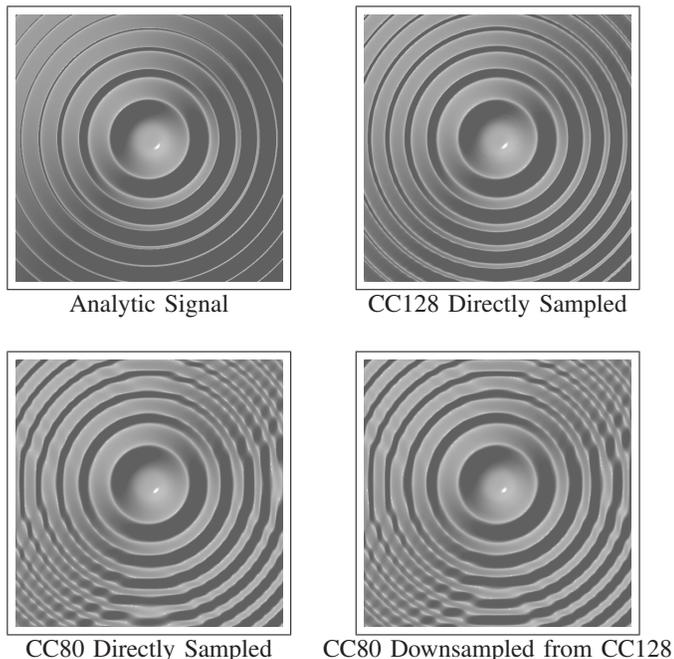


Fig. 6. Downsampling vs. direct sampling of the ML signal.

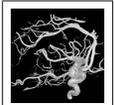
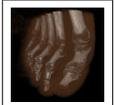
B. Checking for Sampling Resolution Confounds

Our perceptual studies estimated the performance of BCC and FCC samples against three fixed CC sampling resolutions. The choice of CC sampling resolutions may have been a confound. In this section, we use numerical comparability to estimate the relative perceptual behavior of CC, BCC, and FCC samples across a broader range of CC resolutions.

To evaluate the L_2 errors of a range of sampling resolutions of the ML and Fish Tail signals, both were sampled at a range of CC resolutions from $80 \times 80 \times 80$ to $180 \times 180 \times 180$. The signals were also sampled at BCC and FCC resolutions from approximately $0.7 \times 80 \times 80 \times 80$ to $0.7 \times 180 \times 180 \times 180$.

Figure 7 (left column) plots the mean L_2 errors against the number of samples. For every pair of reconstruction filters, the relative ratios of the number of samples required to achieve the same L_2 errors are plotted in Figure 7 (right column). The ratios between the error curves of any two reconstruction filters are roughly constant. We estimate the ratio for each pair of L_2 curves by averaging the values of the corresponding ratio curve.

TABLE II
GENERALIZATIONS TO ADDITIONAL SIGNALS BASED ON NUMERICAL
COMPARABILITY.

Signal	Truth Res	Downsampled Res	Numerical Comparability
Fish Tail 	CC256	CC140 vs. BCC CC140 vs. FCC	69% – 71% 69% – 71%
		CC180 vs. BCC CC180 vs. FCC	69% – 70% 69% – 71%
Skull 	CC256	CC140 vs. BCC CC140 vs. FCC	71% – 74% 69% – 71%
		CC180 vs. BCC CC180 vs. FCC	74% – 76% 73% – 75%
Aneurism 	CC256	CC140 vs. BCC CC140 vs. FCC	77% – 79% 77% – 79%
		CC180 vs. BCC CC180 vs. FCC	79% – 81% 79% – 82%
Foot 	CC256	CC140 vs. BCC CC140 vs. FCC	77% – 79% 74% – 76%
		CC180 vs. BCC CC180 vs. FCC	79% – 80% 76% – 78%

The BCC to CC error ratio and the FCC to CC error ratio are approximately 70% for ML and 72% for Fish Tail. Given the observed correlation between numerical and visual comparability for the images in the perceptual study, these error ratios suggest that the BCC/FCC visual ranges of comparability will also have midpoints close to 70% across the sampling resolutions in Figure 7.

C. Generalizing Beyond The Two Signals

Signal choice is another potential limitation to our perceptual study. We use numerical comparability to estimate the visual comparability of three new datasets, Skull, Aneurism, and Foot, previously used in the biomedical visualization community for evaluation of various reconstruction schemes. The downsampling procedure described in Section III-D and the reconstruction filters described in Section III-H were used to estimate the range of numerical comparability for these signals. The BCC and FCC lattices typically achieve numerical comparability with the CC lattice using 20% – 30% fewer samples (Table II). Though these savings are lower than the 30% – 35% saving for visually comparable results in our perceptual study, they nonetheless suggest that for these three additional signals, the BCC and FCC lattices can provide a visually comparable result using fewer samples than the CC lattice.

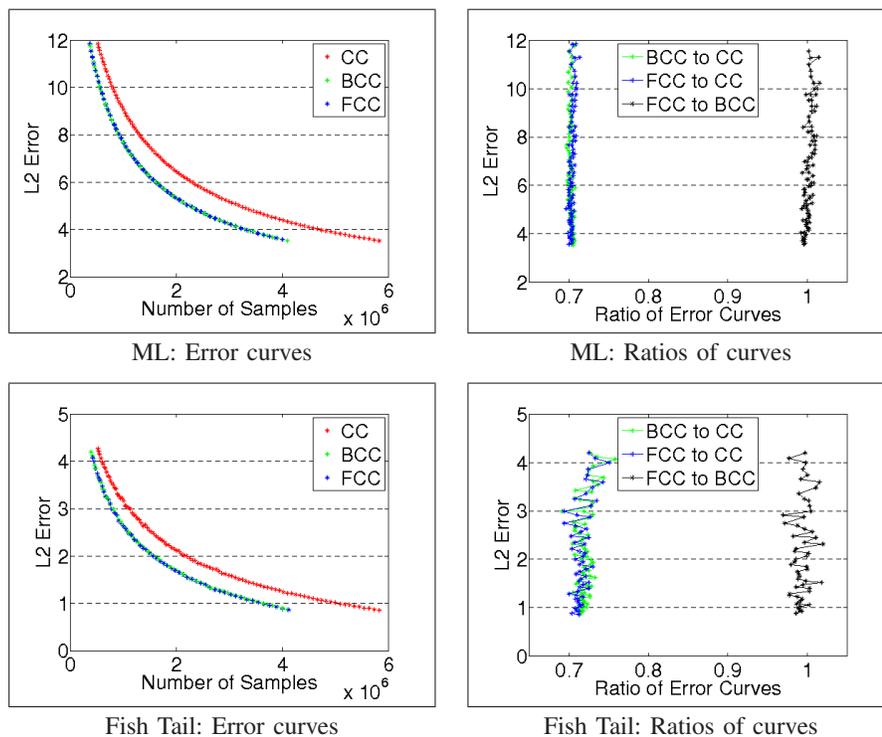


Fig. 7. L_2 errors of a wide range of CC, BCC, and FCC sampling resolutions. Left column: Errors. Right column: Ratios of errors.

D. The Potential Impact of Other Parameter Choices

The parameter space surrounding our perceptual study is enormous, and the study only explores a tiny portion. The analysis of numerical comparability in the previous sections suggests that the perceptual results apply to a range of the space beyond what was directly tested, but that extended range is still only a fraction of the total space. In this section, we argue that the perceptual results are likely to apply even more broadly, across the entire space. We consider it highly unlikely that changing the values of most parameters would change the relative performance of BCC, FCC, and CC.

There are eight external parameters to our study: the signal, the reconstruction filter, the downsampling technique (for signals provided as samples, such as Fish Tail), the camera angle and location, and the rendering choices of transfer function, lighting direction, and shading. This parameter space is large enough that there might be combinations of values that generate images where the BCC and FCC versions require as many or more samples as CC to produce visually comparable results, contradicting our results.

For this to occur, however, the key question is whether varying a parameter will affect the perceptual comparability of BCC or FCC *differentially*, or whether the change will affect all images generated from all three sampling lattices *in the same way*. For most of the parameters to our perceptual study, differential effects appear unlikely. We consider each parameter in turn.

The range of possible signals is vast. The perceptual study results hold for two signals with distinct characteristics. The numerical comparability results suggest these results hold for three more signals. For the results to fail for other signals, there would have to be some property shared by ML and Fish Tail that, when sampled by BCC and FCC and rendered according to our parameter choices, makes the resulting images visually

comparable to images generated from a higher resolution of CC, *and* this property shared by ML and Fish Tail would have to be absent from most other signals. The likelihood of all these conditions holding seems low. Nonetheless, the single most important extension of our perceptual study would be to repeat the protocol with additional signals.

For the reconstruction filters, there are no “good” alternatives “far away” from the ones we selected. That is, we picked filters according to current best practice for each lattice.

There are alternative downsampling approaches to the one we applied to Fish Tail. It is difficult to estimate the degree to which our downsampling method might have differentially affected the images from each lattice. The numerical comparability results, together with visual inspections, suggest there is little to no differential effect, but this possibility should be explored by using at least one other downsampling technique in future studies.

The camera angle and location have mild potential for differential effects. For ML, we believe the two angles of view and single distance used in our perceptual study are sufficient. For Fish Tail, however, there is a real possibility that viewing a different region of the signal, or from a different angle, might change the visual comparability of BCC/FCC images and CC images. The best approach to broadening the range of camera viewpoints, though, is in conjunction with perceptual studies that use different signals.

The rendering choices of transfer function, lighting direction, and shading algorithm appear to have no specific bias favoring images generated by BCC and FCC. It is unlikely that different choices for these parameters would differentially affect some lattices in the perceptual study.

VI. CONCLUSIONS

The BCC and FCC sampling lattices support images of comparable quality to the CC lattice, using 30% – 35% fewer samples. Two forms of evidence support this conclusion. First, a total of 24 observers found the sparser BCC and FCC samples to have comparable visual quality to the denser CC samples for images rendered from two signals. Second, the numerical L_2 errors between a signal reconstructed from the BCC or FCC sampling and the original signal are comparable to the L_2 errors computed from a denser CC sample of the same signal. The ranges of this numerical comparability match the range of visual comparability found by the observers in the perceptual study. The evidence from the L_2 analysis is provisional: Although prior work [17] has argued for the limited usefulness of numeric comparisons as predictors of perceptual comparability, we have found a high correlation for our images and viewing conditions. More work is necessary to understand the relation between numeric and perceptual comparability.

The range of signals, sampling resolutions, reconstruction filters, and camera views examined in this study was relatively small. The numerical comparability analysis suggests that the perceptual results generalize across a broader range, but the ultimate test will be further perceptual studies with different parameter values. While acknowledging the range of values left unconsidered by our study, we also emphasize that there are strong *a priori* reasons to believe our results will generalize across many parameters. Of the alternative parameter values that might be tested in future studies, using different signals appears most important, followed by using at least one other downsampling technique. Future studies could use data sets from other application areas and also develop user tasks that more directly correspond to actual uses of volumetric data, such as medical diagnosis.

Combining this study and previous theoretical and algorithmic results [4], [5], [7], we now have evidence that sampling and reconstruction on the BCC and FCC lattices are more accurate, computationally more efficient, and yet produce visually comparable images to sampling and reconstruction on the popular CC lattice. If our perceptual results are confirmed for a variety of other signals, it will make a strong case for BCC and FCC sampling for 3D visualization and rendering.

ACKNOWLEDGMENTS

We thank Steven Bergner for automating the rendering pipeline, and Leila Kalantari for co-administering the user studies. This research has been supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Canada Foundation for Innovation (CFI). We acknowledge the following for the datasets: Dr. Christof Rezk-Salama and Dr. Michael Scheuering for Fish Tail; Siemens Medical Solutions, Forchheim, Germany for Skull; Philips Research, Hamburg, Germany for Aneurism and Foot; and Dr. Dirk Bartz for making the Skull, Aneurism, and Foot datasets publicly available on <http://www.volvis.org>.

REFERENCES

- [1] S. Daly. In A. B. Watson, editor, *The Visible Differences Predictor: an Algorithm for the Assessment of Image Fidelity*, Digital Images and Human Vision, pages 179–206. MIT Press, Cambridge, MA, 1993.
- [2] B. Efron. Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, 82(397):171–185, 1987.
- [3] K. Engel, M. Hadwiger, J. M. Kniss, C. Rezk-Salama, and D. Weiskopf. *Real-Time Volume Graphics*. A K Peters, Wellesley, MA, 2006.
- [4] A. Entezari. *Optimal Sampling Lattices and Trivariate Box Splines*. PhD thesis, Simon Fraser University, Canada, July 2007.
- [5] A. Entezari, R. Dyer, and T. Möller. Linear and Cubic Box Splines for the Body Centered Cubic Lattice. In *Proceedings of IEEE Visualization 2004*, pages 11–18, 2004.
- [6] A. Entezari, T. Meng, S. Bergner, and T. Möller. A Granular Three Dimensional Multiresolution Transform. In *Proceedings of Eurographics / IEEE VGTC Symposium on Visualization 2006 (EuroVis 2006)*, pages 267–274, 2006.
- [7] A. Entezari, D. Van De Ville, and T. Möller. Practical Box Splines for Volume Rendering on the Body Centered Cubic Lattice. *IEEE Transactions on Visualization and Computer Graphics*, 14(2):313–328, 2008.
- [8] G. Farin. *Curves and Surfaces for Computer-Aided Geometric Design*. Academic Press Inc., San Diego, CA, 4th edition, 1997.
- [9] A. Gaddipatti, R. Machiraju, and R. Yagel. Steering Image Generation with Wavelet Based Perceptual Metric. *Computer Graphics Forum*, 16(3):241–252, 1997.
- [10] A. Kaufman and K. Mueller. Overview of Volume Rendering. In C. D. Hansen and C. R. Johnson, editors, *The Visualization Handbook*, pages 127–174. Elsevier-Butterworth Heinemann: Amsterdam; Boston, MA, 2005.
- [11] J. Kovačević and M. Vetterli. FCO Sampling of Digital Video Using Perfect Reconstruction Filter Banks. *IEEE Transactions on Image Processing*, 2(1):118–122, January 1993.
- [12] H. R. Künsch, E. Agrell, and F. A. Hamprecht. Optimal Lattices for Sampling. *IEEE Transactions on Information Theory*, 51(2):634–47, 2005.
- [13] J. Lubin. In A. B. Watson, editor, *The Use of Psychophysical Data and Models in the Analysis of Display System Performance*, Digital Images and Human Vision, pages 163–178. MIT Press, Cambridge, MA, 1993.
- [14] S. R. Marschner and R. J. Lobb. An Evaluation of Reconstruction Filters for Volume Rendering. In *Proceedings of IEEE Visualization 1994*, pages 100–107, 1994.
- [15] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. Perceptual Blur and Ringing Metrics: Application to JPEG2000. *Signal Processing: Image Communication*, 19:163–172, 2004.
- [16] T. Meng, B. Smith, A. Entezari, A. E. Kirkpatrick, D. Weiskopf, L. Kalantari, and T. Möller. On Visual Quality of Optimal 3D Sampling and Reconstruction. In *Proceedings of Graphics Interface 2007*, pages 265–272, 2007.
- [17] D. P. Mitchell and A. N. Netravali. Reconstruction Filters in Computer Graphics. *Computer Graphics (ACM SIGGRAPH 1988 Proceedings)*, 22(4):221–228, 1988.
- [18] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck. *Discrete-Time Signal Processing*. Prentice Hall: Upper Saddle River, NJ, 2nd edition, 1999.
- [19] D. P. Petersen and D. Middleton. Sampling and Reconstruction of Wave-Number-Limited Functions in N -Dimensional Euclidean Spaces. *Information and Control*, 5(4):279–323, Dec. 1962.
- [20] D. F. Rogers. *An Introduction to NURBS : With Historical Perspective*. Morgan Kaufmann Publishers, San Francisco, CA, 2001.
- [21] F. A. Wichmann and N. J. Hill. The Psychometric Function: I. Fitting, Sampling, and Goodness of Fit. *Perception & Psychophysics*, 8(63):1293–1313, 2001.
- [22] F. A. Wichmann and N. J. Hill. The Psychometric Function: II. Bootstrap-Based Confidence Intervals and Sampling. *Perception & Psychophysics*, 8(63):1314–1329, 2001.

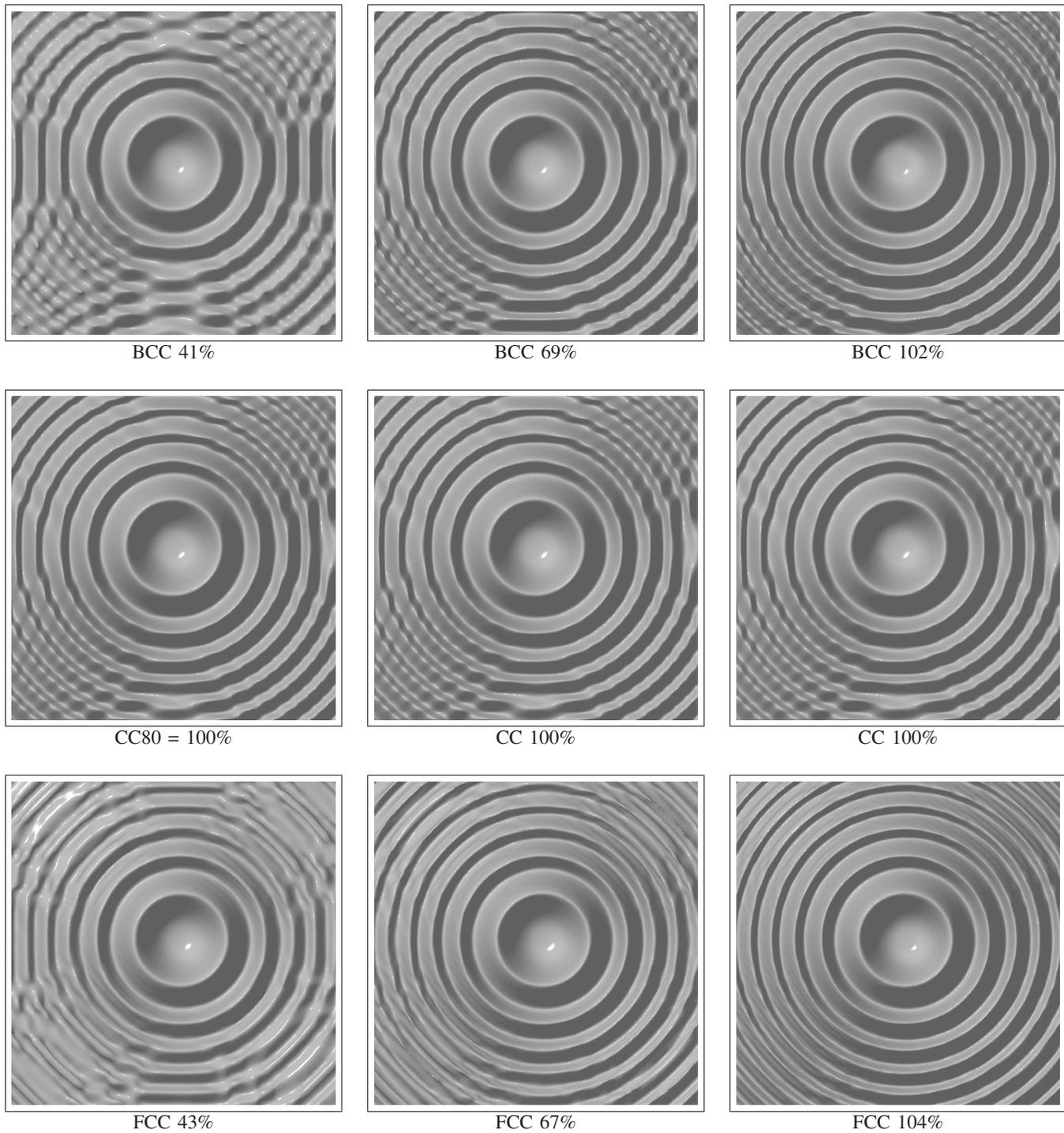


Fig. 8. Images of ML sampled data. Sampling resolutions are shown relative to the fixed CC resolution, which is labeled as 100%. The CC image is duplicated three times for ease of comparison. The images offer confirming evidence that BCC and FCC 65% – 70% (relative to CC) is where comparable visual fidelity occurs for ML (compare the middle columns to the other columns). Note that the diagonal asymmetry in the ML images is due to the lighting direction; if the lighting direction were straight-on, the two diagonals would appear symmetric.

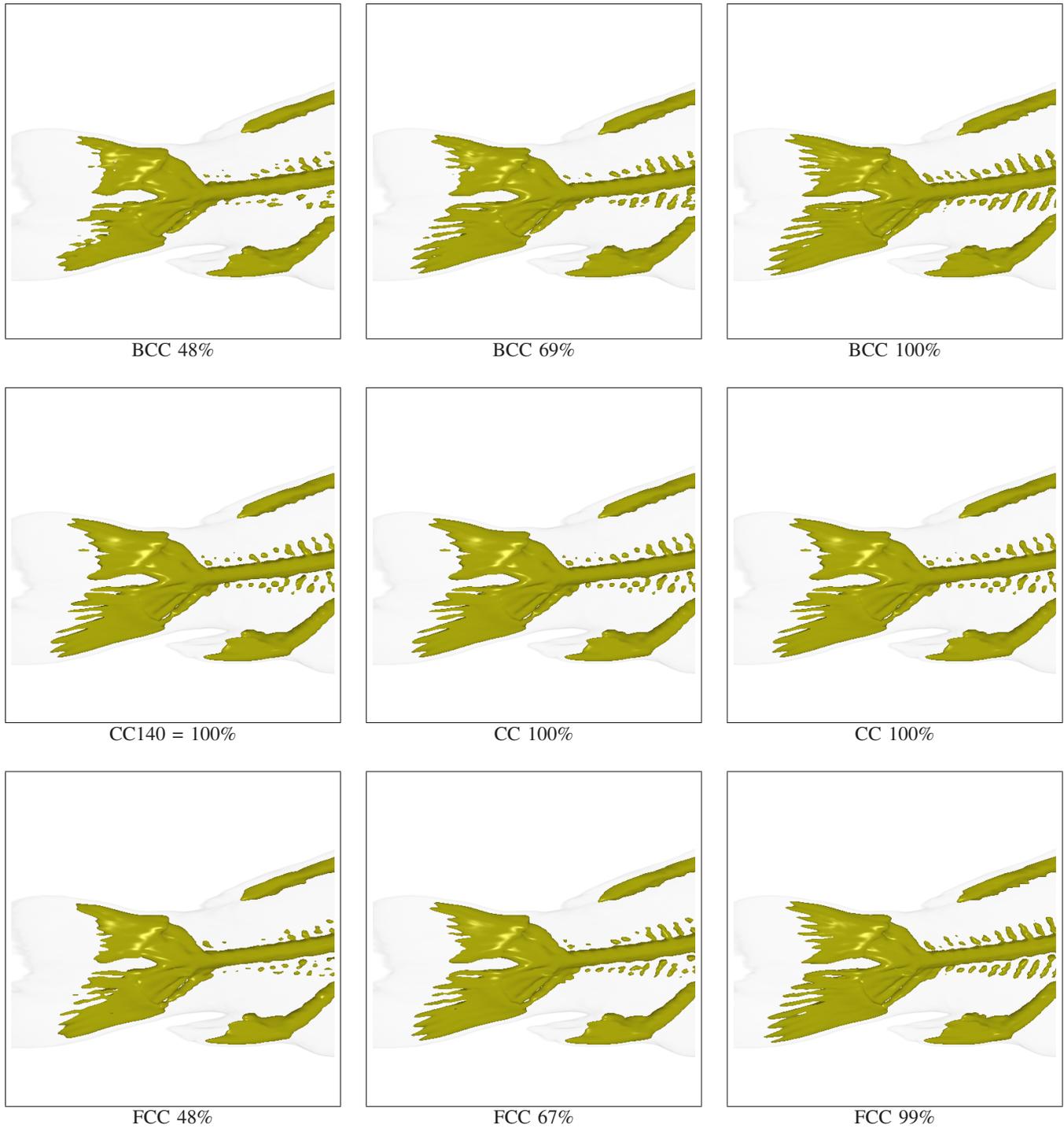


Fig. 9. Images of Fish Tail sampled data. Sampling resolutions are shown relative to the fixed CC resolution, which is labeled as 100%. The CC image is duplicated three times for ease of comparison. The images offer confirming evidence that BCC and FCC 65% – 70% (relative to CC) is where comparable visual fidelity occurs for Fish Tail (compare the middle columns to the other columns).



Tai Meng completed his MSc. in Computing Science at Simon Fraser University in 2008. In collaboration with fellow members of the Graphics, Usability and Visualization (GrUVi) lab, his research work focuses on the visual quality of 3D regular sampling and reconstruction.



Daniel Weiskopf received the MSc (Diplom) degree in physics and the PhD degree in physics, both from Eberhard-Karls-Universität Tübingen, Germany, and he received the Habilitation degree in computer science at Universität Stuttgart, Germany. From 2005 to 2007, Dr. Weiskopf was an assistant professor of computing science at Simon Fraser University, Canada. Since 2007, he has been a professor of computer science at the Visualization Research Center, Universität Stuttgart (VISUS) and at the Visualization and Interactive Systems Institute (VIS), Universität Stuttgart. His research interests include scientific visualization, GPU methods, real-time computer graphics, mixed realities, ubiquitous visualization, perception-oriented computer graphics, and special and general relativity. He is member of ACM SIGGRAPH, the Gesellschaft für Informatik, and the IEEE Computer Society.



Alireza Entezari (M'09) received a PhD from the school of computing science at Simon Fraser University, Vancouver, Canada in 2007. He's currently an Assistant Professor at the Computer and Information Science and Engineering department at the University of Florida. He is co-organizing a workshop on sampling and reconstruction at Banff International Research Station in 2010. His research interests include multidimensional signal processing and multivariate splines with particular applications in scientific visualization and biomedical imaging.



Benjamin Smith is a MSc. student in Computing Science at Simon Fraser University. He earned his B.Math in Computer Science the University of Waterloo in 2005. As a member of the Graphics, Usability and Visualization (GrUVi) and Medical Image Analysis (MIAL) labs, his research focus is on the segmentation and analysis of time varying medical data, such as dynamic PET.



Torsten Möller is an associate professor at the School of Computing Science at Simon Fraser University. He received his PhD in Computer and Information Science from Ohio State University in 1999 and a Vordiplom (BSc) in mathematical computer science from Humboldt University of Berlin, Germany. He is a member of IEEE, ACM, and Eurographics. His research interests include the fields of Visualization and Computer Graphics, especially the mathematical foundations thereof.

He is co-director of the Graphics, Usability and Visualization Lab (GrUVi) and serves on the Board of Advisors for the Centre for Scientific Computing at Simon Fraser University. He is the appointed Vice Chair for Publications of the IEEE Visualization and Graphics Technical Committee (VGTC). He has served on a number of program committees and has been papers co-chair for IEEE Visualization, EuroVis, Graphics Interface, and the Workshop on Volume Graphics as well as the Visualization track of the 2007 International Symposium on Visual Computing. He has also co-organized the 2004 Workshop on Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration at the Banff International Research Station, Canada. He is currently serving on the steering committee of the Symposium on Volume Graphics. Further, he is an associate editor for the IEEE Transactions on Visualization and Computer Graphics (TVCG) as well as the Computer Graphics Forum. He is the recipient of the NSERC DAS award.



Arthur E. Kirkpatrick Arthur Kirkpatrick received his Ph.D. in Computing Science from the University of Oregon in 2000. Since 2001 he has been a member of the Computing Science Department at Simon Fraser University, where he holds the rank of Associate Professor. His research interests include interaction techniques and undergraduate computer science education. He is a member of ACM and ACM SIGCHI.