IEEE MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE

E-LETTER

Vol. 4, No. 4, May 2009

CONTENTS

IEEE COMMUNICATIONS SOCIETY

Message from Editor-in-Chief

The MMTC meeting is one month away, we hope most of our members can make it in the beautiful Germany city Dresden, so we get chance to meet with each other and exchange ideas for future cooperation. The MMTC Chair, Dr. Qian Zhang has published the **MMTC meeting agenda** in the next page (page 3) of the same issue, please go over it and let her know if you have any new proposals or issues that would like to be discussed in the meeting.

In this issue, we have eight scientific articles published in the technology session, five of which are concerning about the quality issues in multimedia communications. Indeed how to fairly measure quality and assure satisfactory content quality is always one of the most interesting topics for this community. The set of papers begins with a Distinguished Position paper, Meditations on Visual Quality, delivered by Dr. Alan C. Bovik (University of Texas at Austin, USA). In this article, Dr. Bovik shares his vision and perspective of the current status in the understanding and definition of visual quality, and throws lights on the future trends and promising research directions for quality assessment.

The second paper by Dr. Weisi Lin (Nanyang Technological University, Singapore) is focused on the Just-noticeable difference (JND) concept and its applications to image and video coding and processing.

The third paper by Dr. Lina Karam (Arizona State University, USA) focuses on the objective quality measurement for video communications, where the current development status of various types of quality metrics is overviewed.

In the fourth paper, Drs. Gokce Dane and Khaled El-Maleh (Qualcomm, USA) approach the video quality enhancement efforts from both perspective of encoder and decoder, bit allocation and post-processing; In the article, a few interesting questions and potential directions for further investigations are proposed for our audiences.

In the fifth paper, Drs. Hongfei Du, Jiangchuan Liu, and Jie Liang (Simon Fraser University, Canada) address the quality of service (QoS) issues for the integrated terrestrial-satellite multimedia systems and relevant optimization techniques. In addition, the paper highlights a few key challenges in the end-to-end QoS assurance in satellite system, such as the reception conditions and return link diversities, for further research and investigations.

In this issue, the selected paper for recommendations by our Column Editor, Dr. Chonggang Wang (NEC Laboratories America, USA), is an article just published in the IEEE Infocom 2009 a week ago, studying the bit error patterns in the wireless local area networks.

In the focused technology column, Dr. Antonios Argyriou (Phillips Research, Netherlands) discusses about the signal processing in distributed and embedded wireless networks. In the IG column, Drs. Jiang (Linda) Xie (University of North Carolina-Charlotte, USA) and Xiaoyuan Gu (MIT, USA) point out the challenges of seamless mobility in wireless networks, and the provision of seamless mobility to mobile multimedia services.

At the end, I sincerely invite you to participate into our E-Letter efforts either as a Guest Editor, an author, or an active reader. Please do not hesitate to write to us (haohong@ieee.org) any time to express your opinions or voluntary interest.

As always, I thank all Editors of the E-Letter, and our authors to make this issue successful.

Thank you very much.

Haohong Wang
Editor-in-Chief, MMTC E-Letter

HIGHLIGHT NEWS & INFORMATION

MMTC Meeting Agenda

June 14-18, 2009
Dresden, Germany

Dear all the MMTC members,

It is quite excited that we will have another MMTC meeting coming soon in ICC 2009 at Dresden, Germany from June 14-18. I am looking forward to seeing all of you there for our MMTC meeting, which has a draft agenda as follows.

0. Informal discussion and networking time
1. welcome new members /introduction
2.Last meeting minutes approval (Globecom 2008)
3. MMTC Best Paper Award 2009 winner announcement
4. Report for the recent change in our TC (sub-committee change)
5. Report on GITC initiatives
6. Report on Conferences activities (who will report for each conference?)
   CCNC 2009
   ICC 2009
   ICME 2009
   Globecom 2009
   ICC 2010
   Globecom 2010
7. TAC Report
8. MMTC IGs Reports - all IG chairs
9. Sub-committees Report
10. Publication Report (e.g., activities in terms of special issues on IEEE journals/Magazines)
11. Report for News Letter activity
12. Suggestions & discussions – everyone
13. Adjourn

Our MMTC established its Best Paper Award to recognize outstanding work in the field of Multimedia Communications (please, see http://www.comsoc.org/~mmc/awards.asp). It is my great pleasure to inform you that under Dr. Dapeng Wu's great leadership, our award sub-committee members have spent tremendous efforts on selecting the following two papers for the MMTC best paper award in 2009.

1. Mea Wang, Baochun Li. "R2: Random Push with Random Network Coding in Live Peer-to-Peer Streaming," in IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Peer-to-Peer Streaming Systems, Vol. 25, No. 9, pp. 1655-1666, December 2007.

2. B. Li, S.-S. Xie, G. Y. Keung, J.-C. Liu, I. Stoica, H. Zhang and X.-Y. Zhang, "An Empirical Study of the Coolstreaming+ System," IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Peer-to-Peer Streaming System, 25(9):1627-1639, December 2007.

Please join me to congratulate for the authors of those two papers for this well-deserved award.
As usual we will have a brief ceremony to give the award plaque during one of our meetings held either at IEEE ICC or IEEE Globecom. This year, the authors of the first paper will receive their award plaque in ICC and the authors of the 2nd paper will receive their plaque in Globecom. Please join the MMTC meeting to share this good news.

Again, looking forward to seeing you in Dresden, Germany soon.

Cheers,

Qian Zhang
IEEE MMTC Chair

TECHNOLOGY ADVANCES

Distinguished Position Paper Series

Meditations on Visual Quality

Alan C. Bovik (IEEE Fellow), University of Texas at Austin, USA
bovik@ece.utexas.edu

Can We Define Visual Quality?

During the 2008 pre-election period of two-day political "flaps," a photograph of Vice-Presidential aspirant Sarah Palin appeared on the cover of Newsweek magazine that garnered an extraordinary degree of media attention [1]. Apparently, her supporters felt the cover photograph to be of "too high quality," as it accurately captured minute facial characteristics, including pores, moles, hairs, and other flaws, that ostensibly should not have been revealed. I recall watching a young, highly partisan pundit on CNN (on whom I could visually detect no such flaws), expostulating about the ill treatment accorded Palin, specifically, publishing a picture having high spatial resolution, no photographic touch-ups, and no softening of facial features by image processing methods. The pundit expressed, with considerable fervor, that this was no less than iniquitous. Yet, given that the image was an extreme close-up, no doubt intended to elevate feelings of propinquity in the reader towards Palin, it seemed difficult to me to avoid some degree of realism in the photographic portrayal. Needless to say, while I did not quite view the tirade as an indictment of our collective years of trying to improve the appearances of images, it did re-engender mixed feelings I have about the meaning of the term "quality" as it pertains to perceptual signals, and how this term should be defined, interpreted, and understood.

Indeed, despite the years that my colleagues and I have devoted to the topic of "image quality" and "video quality," I find myself consternated by the fact that many of the terms remain poorly defined, even in cases where it seems that precision in capturing the definiendum has been reached. By way of simultaneous introduction to the topic, I offer some examples.

The problem that has received the most attention has been dubbed No Reference Quality Assessment (of images or videos – or speech, for that matter), or NR QA, for short. The problem to be accomplished is to determine the "quality" of a "test" signal *relative* to a known and presumably high-quality "reference" signal, via an algorithm, and in such a manner that it agrees with the subjective judgments of a sufficiently large ensemble of human subjects that the degree of agreement may be demonstrated with statistical certitude. The comparison with human subjectivity is only important, of course, if the objective quality judgment made by algorithm is intended to predict human judgment; so I assume this here.

Yet, the term "quality" remains imprecise, since after all, the "reference" image being compared to may be of low visual appeal or be distorted. Indeed, in my 30 years of work in this area, I have yet to see a truly "pristine" digitized image – meaning, one that I cannot find any error or flaw with, and that I cannot distinguish from experiential reality. All so-called "reference" images and videos suffer from visible flaws. Further, there exists no algorithm or theory that can determine whether a visual signal is, indeed, "pristine," or if not, how far from "pristine" it should be judged. Therefore, any algorithmic, objective measure of *comparative* image quality must fall short, with the missing ingredient being a statement of the quality of the "reference" image.

Other terms that come to mind (and have been used) are "similarity" and "fidelity." "Similarity" is partially apropos, since most "image quality" indices measure just that, and indeed many have symmetric definitions (yielding the same result if "reference" and "test" signals are reversed), including our own currently fashionable "SSIM index" [2]-[4], [3]. Yet, "similarity" and "quality" are certainly not synonymous, and only become so if the "test" signal is demonstrably "pristine," and if the algorithm agrees with humans that are also striving to assess "quality."

"Fidelity" is a hoary term in our profession, dating from the description of early 20th-century phonograph players as (inevitably) "hi-fi." Yet, perhaps we should blow away any accrued dust, since "fidelity" implies both an *intrinsic* level of presumed signal "goodness," and at the same time, a faithful maintenance of that "goodness." Given that a "reference" signal is a "good" one – an accurate representation of reality – then decreeing that a "test" signal that agrees closely with it is consequently of "high fidelity" is satisfying. Yet even this definition falls short, since the equally hoary mean-squared error passes this test, yet fails miserably when cast against human subjectivity – the Final Arbiter. Perhaps appending a modifier, for example, "visual fidelity," or "perceptual fidelity," solves the problem, as we attempted in naming one of our successful algorithms [5], since the implied presumption is then that the Final Arbiter will cast the ballot. Yet these terms remain imprecise as well, since visual perception hardly defines photo-optical ground truth.

So, why not embrace "quality" and be done with it? Well – I think the term has too many overtone; it correlates with such indefinable concepts as "attractiveness," "aesthetics," and "beauty," as well as accuracy in reproduction. These are topics on which very little foundational work has been done, although there has been some developing interest in these concepts [6].

On reflecting further, I find that my disquietude on this aspect of the topic has increased, as it appears that the answers may lie not only in deeper realms of perception, where I feel little discomfort, but also in the realm of philosophy, which lies outside my topical purlieu. I am reminded of John Ruskin's observation that *Nothing can be beautiful which is not true…* and indeed, after years of contemplating this issue, defining visual Quality still seems as elusive as defining Absolute Truth. And so, I will fall back on half of the ancient aphorism: He who knows, knows not….

A related matter on which I feel solid engineering ground, and can perhaps more concretely contribute is with regards to the widespread use of the term "metric" in defining (what I prefer to call) "index." A *metric* is a precisely defined distance function that is subject to specific conditions, which nearly all "quality metrics" fail to satisfy (excepting the MSE, which is a dreadful image quality index [7]), including such algorithms as Pinson and Wolf's "Video Quality Metric" (VQM) [8], which is an excellent video quality index, but hardly a metric. I can say little else on this, other than apologizing for my own careless use of the term on occasion in the past, and promising an expressed determination to lobby to correct this misuse in the future.

Benchmarking Using QA Algorithms is a Good Idea

The reader should *not* take the foregoing ruminations to suggest that image and video "quality" algorithms, lacking entirely comprehensive definition, are not useful. Actually, they are much more *useful* than they are being *used*. My point in the preceding is that our understanding of the problem remains incomplete, and that there remains considerable room for improvement; yet modern QA algorithms are significantly more efficacious than the most widely-used approaches. So, to the contrary; my hope is that, by reading this piece, multimedia engineers will reconsider their use of traditional quality indices such as the MSE, the PSNR, or packet-loss rate – at least in isolation. These venerable measures correlate poorly with subjective impressions of video quality, despite their other attractive features, notably low complexity.

Instead, I urge the adoption of modern video QA algorithms such as the successful ISO standard VQM [8], some flavor of the effective and efficient Structural SIMilarity (SSIM) index [3], [4], (in particular the multi-scale SSIM index [9]), or the motion-tuned MOVIE index [10]. The state-of-the-art of video QA has reached the point where the disparity in performance between older metrics, such as the MSE or PSNR, and modern VQA algorithms has become quite spacious.

The way in which I think that indices for image, video (and other perceptual signal QA) need to be used is for benchmarking the many flavors of processing algorithms. What is the best video compression algorithm? Or the image restoration device that delivers the most visually appealing results? Amongst the many error resilience schemes developed to protect against packet losses – is there one that best protects the visual quality of the video? What about something as elemental as image denoising – is there a best algorithm? Is it some flavor of wavelet soft thresholding algorithm? Or, perhaps, an anisotropic diffusion approach?

We do not have definitive answers to any of these questions. Of course, analytical arguments can be posited for each competing approach to a processing task – there are nearly as many "optimality" criteria as there are algorithms.

Usually, the MSE or PSNR are quoted, and one or two visual examples offered to the reader. In the end, the only certain method to establish degrees of performance between the multitude of algorithms – (those that ostensibly seek to either improve or not vitiate perceptual quality) – is by conducting large-scale human studies. This is generally infeasible, of course, not only owing to the effort implied – which is substantial, but also to the obsolescence of each study once the technology advances. Ordinarily, such approaches are limited to standards bodies, and even so, the results usually have a short half-life.

Yet I maintain that perceptual QA indices have advanced to the point that their perceptual relevance is sufficiently high and their performance adequately general, that they be regarded as standards of comparison whenever a reference is available. Indeed, it would be satisfying to see broad comparative studies of all the competitive members of each algorithm category, cast against (if not human subjective judgments) perceptual QA scores. Such approaches are being embraced in the computer vision community, as for example, the Middlebury Computer Vision Page [15], which maintains up-to-date comparative studies on computed stereo, optical flow, and other vision algorithms. Cannot our multimedia community do the same?

Naturally, I do not expect my avuncular plea to be hastily answered, since assembling the diverse algorithms is difficult and the effort considerable. Nevertheless, I *do* challenge the community to compare the performance of new algorithms against older ones using perceptually meaningful QA indices. Until the day comes that models of visual brain processing are sufficiently detailed that optimization algorithms can be married to them, perceptual QA algorithms represent our best benchmarking method.

Along a similar vein, most visual signal processing algorithms have been *designed* using optimality criteria of questionable perceptual relevance, whereas better results might be obtained using QA indices as objective functions. Given that little work has been done on this topic, I will save it for a later soliloquy.

Problems that Need to be Solved

Video QA can be viewed as assessing the quality of a process of visual communications; as such, better models of the signal transmitter, the signal receiver, and the overall channel are likely to provide fertile ground for improving QA algorithms. Of course, the transmitter (the physical world which emanates light onto optical sensors) is terrifically complex and difficult to model, as is the receiver (the human eye-brain system). Unless the possible distortion(s) are known, or assumed, and well-modeled, then the channel (anything that modifies the signal between transmission and reception) also is difficult to manage, since video distortions are quite diverse. From this perspective, it becomes clear that the greatest future strides in QA algorithm development will come from better modeling of the transmitter, the receiver, and the channel.

Transmitter Models

Video transmitter modeling is complex since the interaction of light with the world is extraordinarily complex, owing to the infinite variety of surface shapes, reflectance profiles, and incident irradiance patterns. Yet the visual world can be broadly described as mostly smooth at the scale of observation, broken by relatively sparse discontinuities and irregularities, and exhibiting scale invariance properties. These observations have been used to develop simple natural image statistical models that have proved quite useful for many image processing tasks [11], including image QA [5] and video QA [10]. I expect that, going forward, further refinement of so-called natural scene statistic (NSS) models, especially with regards to more accurately capturing spatial interactions, will greatly advance visual signal QA, and I view this as a particularly key research direction.

Receiver Models

The dual problem, of course, is modeling the receiver that has evolved over millions of years to optimally extract information from natural scenes. Ideally, this should encompass models not only of optical and retinal processing at the front-end of the vision system, which is well-modeled, but also modeling of intermediate processing in primary (or striate) visual cortex (area V1, moderately well-understood), the lateral geniculate nucleus, or LGN (not well understood, although thought to be involved in temporal de-correlation of visual data), and extra-striate cortical areas, such as MT (area V5), which is implicated in motion perception and eye movements, area V2 processing (possible pre-processing for object recognition), the inferior temporal gyrus (object recognition, apparently), and other regions of the brain. Indeed, if I felt I had one grain of wisdom to impart to researchers on video QA, it would be to acquire a basic understanding of cortical visual processing, as improved algorithms for visual QA and other visual processing tasks will likely rely on improved visual receiver models. A good place to start is a modern overview article authored by several prominent vision scientists, which I think is nicely accessible to multimedia engineers [16].

A good example is our recent use of extra-cortical models of motion processing to improve video QA algorithms. Models of motion perception have been used in the past in this regard, but only by simple "black box" temporal filtering models of contrast sensitivity. For example, VQM [8] operates by examining small spatio-temporal blocks, while other algorithms have deployed one or two temporal filters to capture "motion energy" [10]. Our own recently-developed MOVIE index goes beyond this by decomposing video data using a 3D (area V1 model) Gabor wavelet decomposition [13], the responses of which feed excitatory-inhibitory processing to compute video quality along computed motion trajectories [10], using a simple but effective model of motion-tuned processing in extra-striate area MT [12]. Such an approach leads to dramatically better results than the traditional MSE or PSNR, relative to human subjective judgment. For example; on the soon-to-be-released LIVE Video QA Dataset and Study (see below) [14], which includes MPEG, H.264, IP packet-loss and wireless packet-loss videos, tested on nearly 40 human subjects, the MOVIE index delivers a Spearman Rank Order Correlation Coefficient (SROCC) of 0.76, as compared to a paltry 0.37 for the PSNR. The standardized VQM algorithm also performed quite well, with an SROCC of 0.70, while possessing the advantage of a fast available implementation. VQM is freely downloadable from the NTIA website and can be used by anyone.

Channel Models

Two broad approaches can be pursued for capturing distortions that lead to impaired visual quality. The *specific* approach is to model one or more distortions that can occur in an application of interest, and then seek occurrences that match the distortion model. The *general* approach is to model "normal" video behavior, and then seek departures from that behavior. This applies to all categorizations of QA algorithm, whether Full-, Reduced-, or No-Reference (FR, RR, or NR). I won't talk about the specific approaches and what needs to be done (since there are too many), other than to say that foreknowledge of the distortion(s) is potent information, and that merging specific approaches with the general is most promising.

The general approach is inseparable from the transmitter modeling problem, since the idea is to assess departures from "statistical naturalness." The most general (NR) scenario remains quite difficult, owing to the simplicity of current NSS models. If information on the source statistics of the reference videos can be measured, then a very good RR algorithm can be derived, even if nothing is know about the distortion, as exemplified by Li and Wang's "general-purpose" RR image QA algorithm [17]. If the distortions are known and can be modeled against "normality" of the source statistics, then good NR algorithms can even be derived, as in [18]. If the reference signal is entirely available, then an algorithm that models the information that is lost by distortion, as measured by modification of the source statistics, can perform exceedingly well [5].

Saliency and Visual Attention

Humans possess highly mobile heads, and eyeballs that are freely directed around their orbits, focusing complete visual attention and visual resources only on a small portion of a video at a given moment. Conversely, both image features and distortion features may vary in their attraction, and image and distortion features may mask one another. Moreover, distortions may be localized in space and or time. All of these aspects suggest that the perceptual salience or conspicuity of features – their ability to draw visual fixations – may be integral to understanding and improving visual QA. Indeed, spatial saliency features have been explored in this regard [19], with some benefit. Temporal features are more germane, since motion can be a strong visual attractor, but less has been done on understanding the statistical nature of temporal attractors. There is also the necessity of separating interesting motions (e.g., of moving objects) from less interesting motions (e.g., ego-motion). Nevertheless, I view this as a most interesting direction of inquiry, although to be done properly, more research needs to be done on measuring the statistics of videos at the point of gaze, following similar work on static images [20].

The main difficulty is that while we have a good sense of the low-level features that attract visual attention (isolated changes in contrast, colors, motion, etc.), visual fixations are largely guided by higher-level factors, such as the locations of recognizable objects, visual search goals, aural cues, and so on. This is one of the most stimulating directions of inquiry, not only for QA, but for understanding general visual processes. It is another example of the deeply commingled intertwining of perceptual and in this case behavior science with the field of video engineering. The juxtaposition of these fields is among the most fruitful and promising in both brain and information science, yet relatively few video engineers and vision scientists exert their efforts in both realms.

Other Directions

I haven't mentioned the use of chromatic information in video QA. Naturally, it is a significant and important topic, and certainly relevant to quality perception. Yet, little color-specific research has been done. Indeed, many QA algorithms either do not use chromatic information, or deploy it in a desultory way, viz., simply apply a luminance QA scheme to the chromatic components. Yet, since color quality forms such an important part of our perception of quality ("wrong" colors are certainly quite annoying), I anticipate future efforts on this topic.

The issue of *stereoscopic* (3D) image/video quality is an important emerging one. Currently, graphics card companies are developing stereoscopic 3D gaming platforms, and no doubt, this will soon be the standard environment. Digital 3D cinema is blooming, and more general applications, both consumer and scientific, are certain to follow quickly. The stereo QA problem is greatly complicated by the lack of an available 3D reference image, meaning, a 3D reference scene that the human stereoscopic "cyclopean image" can be compared to. I look forward to clever developments in solving this timely problem.

On Quality Assessment Databases

Quality assessment algorithms only possess value if they can be shown, with statistical certainty, to correlate with human subjective judgments of quality. The *de facto* standard dataset of distortions and human scores for assessing the performance of *still image* QA algorithms is the LIVE Image Quality Database [22], which is freely available for download. The industry organization most associated with such tests for videos, and with associated standardization activities, is the Video Quality Experts Group (VQEG). The VQEG has completed and is still involved in several ongoing studies of video quality.

However, while the VQEG has made available the subjective data and videos from their first (deeply flawed) study – the FRTV Phase 1 study [21] – they have not made available the videos or data from later studies, and have indicated that such data will *not* be made available. Indeed, this is true even of the follow-up VQEG FRTV Phase 2 study (which the algorithm VQM emerged as the clear winner of, leading to its standardization). Only the subjective data was made available publicly while the videos are not available, for a variety of reasons. Correspondence with members of VQEG indicates that there are "copyright and licensing issues" involved with the videos, which prevents them from being made public. Yet, I take issue with an industry-driven organization that deploys source and distorted videos that were supplied by the proponents, that conducts tests in secret, and that reports the results in restricted form. In such as situation, no algorithm can compete with the industry proponents. The only recourse for an algorithm developer would be to make requests from the participating organizations that contributed videos directly. Unfortunately, most of these videos are no longer available, even to VQEG members!

One ostensible reason that is given by VQEG members is that making the data available would enable algorithm developers to "train" their algorithms to perform well on the specific dataset. I reject that argument, for three reasons. First, if a dataset is easy to train on, then I submit that it lacks adequate diversity of both content and distortion (type and degree). Secondly, if an algorithm is reported with good results, then it should be made available as an executable, and its performance relative to parameter variation studied. Good examples are the SSIM and multi-scale SSIM (MS-SSIM) indices, which perform well even when the (few) parameters are varied, and VQM, which performs well on every dataset we have tested it on, although it is a heavily trained algorithm [8].

The situation with the forthcoming VQEG Multimedia dataset is identical: the VQEG plans to release *only* the subjective data (in September, 2009) and the videos will *not* be publicly released.

For these and other reasons, LIVE is developing publicly-available video QA databases that will supplement the LIVE Image Quality Database. This has been a major undertaking involving the acquisition of a large number of high-quality videos; creating a wide diversity of distorted videos – carefully separated by perceptual levels of distortion – and incorporating MPEG and H.264 compression artifacts as well as simulated IP and wireless channel errors. We have performed extensive human studies and statistical tests on the results of a variety of video QA algorithms on the datasets. We expect to release our results and databases shortly [14], [25].

Summary

I hope that I've been able to stimulate additional interest in what I view as the future of research in quality assessment. Hopefully, some will reconsider using the PSNR as their standard of comparison, and others will be inspired to become not just video engineers, but also "perception engineers." In any case, this is an exciting field that is still young and presents many challenges. Foremost amongst these, in my view, is the "blind" or NR problem since (coming full circle) within the solution to it lies the answer to the question of what, after all, "quality" means.

References

[1] J. Meacham, "She's one of the folks," *Newsweek*, cover photograph, Sept. 2008.
[2] Z. Wang and A.C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.
[3] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.
[4] Z. Wang, L. Lu, and A.C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 121–132, Feb. 2004.
[5] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, 2006.
[6] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Studying aesthetics in photographic images using a computational approach," *Proc. European Conf. Comput. Vision*, Graz (Austria), 2006.
[7] Z. Wang and A.C. Bovik, "Mean squared error: Love it or leave it? — A new look at signal fidelity measures," *IEEE Signal Process Mag*, vol. 26, pp. 98-117, Jan. 2009.
[8] M.H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans on Broadcasting*, vol. 50, pp. 312–322, Sept. 2004.
[9] Z. Wang, E.P. Simoncelli and A.C. Bovik, "Multi-scale structural similarity for image quality assessment," *IEEE Conf Sig, Syst., Comput.*, Asilomar, CA Nov. 2003.
[10] K. Seshadrinathan and A.C. Bovik, "Motion-tuned spatio-temporal quality assessment of natural videos," *IEEE Transactions on Image Processing*, revised (preprint by request).
[11] E. Simoncelli, "Statistical modeling of photographic images," in *The Handbook of Image and Video Processing* (A.C. Bovik, Ed.), pp. 431-441, May 2005.
[12] E.P. Simoncelli and D.J. Heeger, "A model of neuronal responses in visual area MT," *Vision Research*, vol. 38, no. 5, pp.743–761, Mar 1998.
[13] A.C. Bovik, M. Clark, and W.S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55–73, Jan. 1990.
[14] K. Seshadrinathan, R. Soundararajan and A.C. Bovik, "Study of subjective and objective quality assessment of video," *IEEE Transactions on Image Processing*, submitted (preprint by request).
[15] D. Scharstein and R. Szeliski, 'The Middlebury Computer Vision Pages," [online], Available at: http://vision.middlebury.edu.

[16] M. Carandini, J.B. Demb, V. Mante, D.J. Tolhurst, Y. Dan, B.A. Olshausen, J.L. Gallant, and N.C. Rust, "Do we know what the early visual system does?," *The Journal of Neuroscience*, vol. 25, no. 46, pp. 10577-10597, Nov. 2005.

[17] Q. Li and Z. Wang, "General-purpose reduced-reference image quality assessment based on perceptually and statistically motivated image representation," *IEEE Int'l Conf Image Process*, San Diego, CA, Oct. 2008.

[18] H.R. Sheikh, A.C. Bovik and L.K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans Image Process*, vol. 14, no. 11, pp. 1918-1927, Nov. 2005.

[19] A.K. Moorthy and A.C. Bovik, "Visual importance pooling for image quality assessment," *IEEE J. Special Topics Signal Process*, Special Issue on Visual Media Quality Assessment, to appear (preprint by request).

[20] U. Rajashekar, I. van der Linde, A.C. Bovik and L.K. Cormack, "Foveated analysis of image features at fixations," *Vision Research*, vol. 47, pp. 3160-3172, Nov 2007.

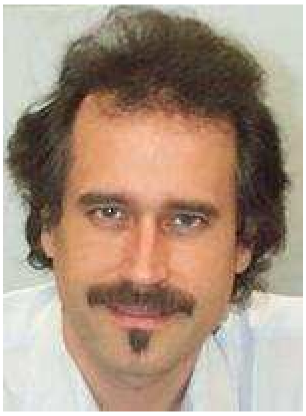
[21] H.R. Sheikh, M.F. Sabir and A.C. Bovik, "An evaluation of recent full reference image quality assessment algorithms," *IEEE Trans Image Process*, vol. 15, no. 11, pp. 3440-3451, Nov 2006.

[22] H.R Sheikh, Z. Wang, L.K. Cormack and A.C. Bovik, *The LIVE Image Quality Database*, September 2005 [Online]. Available: http://live.ece.utexas.edu/research/quality/subjective.htm.

[23] A. Webster, "Progress and future plans for VQEG," in *ETSI STQ Workshop on Multimedia Quality of Service*, 2008. [Online]. Available: http://portal.etsi.org/docbox/Workshop/2008/2008 06 STQWORKSHOP/VQEG ArthurWebster.pdf.

[24] The Video Quality Experts Group. (2000) Final report from the video quality experts group on the validation of objective quality metrics for video quality assessment. [Online]. Available: http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI.

[25] A.K. Moorthy and A.C. Bovik, "Wireless video quality assessment: A study of subjective scores and objective algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, submitted (preprint by request).



Al Bovik is the Curry/Cullen Trust Endowed Chair Professor at The University of Texas at Austin. He is well known for his fundamental work in nonlinear image processing, perceptual image and video processing and computational modeling of visual perception. Al is particularly noted for his pioneering work on robust image processing, multi-resolution image processing, image and video quality assessment, and computational modeling of visual perception. He is also noted for innovations in engineering education, including his popular books and the widely-used Signal, Image and Video Audiovisual (SIVA) Demonstration Gallery. Al has also been a major contributor to engineering service, including innovating and creating the *IEEE International Conference on Image Processing*, first held in Austin, Texas, in November, 1994, and playing an instrumental role in proposing and creating the *IEEE Transactions on Image Processing*, for which he served as Editor-in-Chief for six years. Al has received a number of major awards from the IEEE Signal Processing Society, including: the Education Award (2008); the Technical Achievement Award (2005), the Distinguished Lecturer Award (2000); and the Meritorious Service Award (1998). He is a Fellow of the IEEE, a Fellow of the Optical Society of America (OSA), and a Fellow of the Society of Photo-Optical and Instrumentation Engineers (SPIE).

Determine Visual Just-noticeable Difference (JND) for Multimedia Applications

Weisi Lin, Nanyang Technological University, Singapore
WSLin@ntu.edu.sg

Visual signal is acquired, synthesized, enhanced, watermarked, compressed, transmitted, stored, reconstructed, evaluated, authenticated, displayed, or printed before being presented to the human visual system (HVS). It is well known that the HVS cannot sense all changes in an image due to its underlying physiological and psychological mechanisms, and therefore advantageous to incorporate knowledge of the HVS visibility thresholds into visual processing algorithms/systems, since the HVS is the ultimate receiver of the majority of processed images and video. With perceptual knowledge, the scarce system resources (computing power, bandwidth, memory space, display/printing resolution, and so on) can be allocated to achieve the maximum perceptual significance, accessory information (e.g., for watermarking, authentication, and error protection) can be concealed in the regions with the least HVS sensitivity to the incurred changes, and visual quality of processed images can be evaluated for better alignment with the human perception. Incorporating the HVS visibility thresholds appropriately can play an important role in shaping and optimizing many image processing algorithms.

Just-noticeable difference (JND) refers to the visibility threshold below which any change cannot be detected by the HVS [1-3]. Its determination in general is complex and challenging, because this is related to the HVS characteristics, as well as some recognition process in the human brain, and is adaptive to the contents of the visual signal under consideration. Other affecting factors include viewing conditions (such as viewing distance, ambient lighting, the context of preceding display, the pixel position in the image), and even viewers' preference and experience (e.g., trained observers are able to spot certain changes more easily). The JND is usually defined as the visibility threshold for the majority (e.g., 75%) of ordinary observers under typical viewing conditions.

1. JND with Subbands

In literature, the JND in subbands has been formulated as the product of a base threshold due to the spatial CSF (Contrast Sensitivity Function) and a number of elevation parameters due to other effects. Let *n* denote the position of a DCT block, and (*k*, *l*) denote a DCT subband. The JND can be expressed as:

$$s(n,k,l) = t_{s-csf}(n,k,l)\prod_{\zeta}\alpha_{\zeta}(n,k,l)$$

where $t_{s-csf}(n,k,l)$ is the base threshold due to the spatial CSF [4], and α_{ζ} (*n*, *k*, *l*) is the elevation parameter due to the effect ζ (representing luminance adaptation [5,6], intra-band masking [5,7], inter-band masking [8], color channel masking [9], temporal CSF [10,11], etc.). The JND can be also determined in the wavelet [12] and other subbands [13].

2. JND with Pixels

Pixel-wise JND can be derived from pixel domain [14,15] or from subbands [13,16]. In the former case, luminance adaptation and texture masking are the major factors being considered, while in the latter case, spatial CSF can be also incorporated for more accurate estimation. In the case of images, the pixel-wise JND at the pisition (x,y) can be estimated as [15]:

$$\begin{aligned} S(x,y) = {} & T^{l}(x,y) + T^{t}(x,y) \\ & - c^{lt}(x,y)\cdot\min\{T^{l}(x,y), T^{t}(x,y)\} \end{aligned}$$

where $T^l(x,y)$ and $T^t(x,y)$ are the visibility thresholds for luminance adaptation and texture masking [14], respectively; $c^{lt}(x,y)$ accounts for the overlapping effect in masking, and $0 < c^{lt}(x,y) \leq 1$. For video, temporal (interframe) masking effect can be incorporated [17].

The contrast sensitivity reaches its maximum at the fovea and decreases towards the peripheral retina. The JND model represents the visibility threshold when the HVS attention is there. The overall visual sensitivity at a location in the image should be the JND modulated by the visual attention map [18]. The foveation model proposed in [19] can be used to derive the overall visual sensitivity by modifying the JND at every location with the eccentricity away from the foveation.

3. JND Model Evaluation

The accuracy of pixel-wise JND model can be evaluated by its effectiveness in shaping noise in an image or video frame [14,3]:

$$\hat{I}(x,y) = I(x,y) + q \cdot s^{random}(x,y) \cdot S(x,y)$$

where $I(x,y)$ is the original image, q (> 0) is a parameter to control the noise level, and $s^{random}(x,y)$ takes either +1 or -1 randomly, regarding x and y, to avoid introduction of fixed pattern of changes. A similar process can be used for a subband based JND model.

Perceptual visual quality of the resultant noise-injected images can be compared and evaluated with subjective viewing tests. The resultant mean opinion score (MOS) is regarded as a fair indicator of perceptual quality for each image if a sufficient number of observers are involved. Under a same level of total error energy (e.g., a same MSE or PSNR), the better perceptual quality the noise-injected image/video has, the more accurate the JND model is; alternatively, with a same level of perceptual visual quality, a more accurate JND model is able to shape more noise (i.e., resulting in lower MSE or PSNR) in an image.

4. Applications

Knowledge on JND no doubt can help in designing, shaping and optimizing many image processing algorithms and systems. For visual quality/distortion prediction, a metric can be defined or fine-tuned according to JND [5, 20] for better matching the HVS perception. JND has been used to determine not only the noticeable visual distortion (as in the majority of existing relevant metrics) but also the possibly noticeable visual quality enhancement (against the original image) [20]. A JND-based perceptual metric can be also adopted beyond the quality evaluation purpose (e.g., for image synthesis [13]). The JND profile facilitates perceptual compression for image and video. Since the quantization process is the major cause of coding errors, proper quantization steps can be adaptively chosen according to the JND for a given bandwidth [5,7,14]. For motion estimation, the JND information helps in deciding suitable estimation mode [21] and efficient search process [22]. The JND determination can bring about new insight in many other manipulations for an image and video coding process, like inter-frame replenishment [23], bit allocation, object-based coding, and filtering of motion estimated residues [22] or DCT coefficients [24]. In many practical applications, certain accessory data have to be embedded inside visual signal itself (e.g., for watermarking, authentication, and error protection). With the JND indication, it is possible to insert such data in an image with minimum visual difference [25].

References

[1] N. Jayant, et al, "Signal compression based on models of human perception", Proc. IEEE, 81:1385-1422, Oct. 1993.

[2] B. W. Keelan, *Handbook of Image Quality*, Marcel Dekker Inc, 2002.

[3] W. Lin, *Computational Models for Just-noticeable Difference*, in Digital Video Image Quality and Perceptual Coding, eds. H. R. Wu and K. R. Rao, CRC Press, 2006..

[4] A. J. Ahumada and H. A. Peterson, "Luminance-model-based DCT quantization for color image compression", SPIE Proc. Human Vision, Visual Processing, and Digital Display III, pp.365-374, 1992.

[5] A. B. Watson, "DCTune: A technique for visual optimization of DCT quantization matrices for individual images", Society for Information Display Digest of Technical Papers XXIV, pp.946-949, 1993.

[6] X. Zhang, et al, "Improved Estimation for Just-noticeable Visual Distortion", *Signal Processing*, 85(4):795-808, April 2005.

[7] I. Hontsch, and L. J. Karam, "Adaptive image coding with perceptual distortioncontrol", IEEE Trans. Image Processing, 11(3):214-222, March 2002.

[8] H. Y. Tong and A. N. Venetsanopoulos, "A perceptual model for JPEG applications based on block classification, texture masking, and luminance masking", IEEE Int'l Conf. Image Processing (ICIP), pp.428-432, 1998.

[9] A. J. Ahumada, W. K, Krebs, "Masking in color images", Proc. of SPIE Human Vision and Electronic Imaging VI, vol. 4299, pp. 187-194, January 2001.

[10] B.Watson, et al, "DVQ: A digital video quality metric based on human vision", Journal of Electronic Imaging, 10(1):20-29, January 2001.

[11] Y. Jia, et al, "Estimating Just-Noticeable Distortion for Video", *IEEE Trans. Circuits and Systems for Video Technology*, 16(7):820- 829, July, 2006.

[12] A. B. Watson, et al, "Visibility of wavelet quantization noise", IEEE Trans. Image

Processing, vol. 6, No. 8, pp. 1164-1175, pp.1164-1175, Aug. 1997.
[13] M. Ramasubramanian, et al, "A perceptual based physical error metric for realistic image synthesis", Computer Graphics (SIGGRAPH'99 Conference Proceedings), 33(4), pp. 73-82, August 1999.
[14] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of Just-Noticeable-Distortion Profile", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 5, No. 6, pp.467-476, Dec 1995.
[15] X. Yang, et al, "Motion-compensated Residue Pre-processing in Video Coding Based on Just-noticeable-distortion Profile", *IEEE Trans. Circuits and Systems for Video Technology*, vol.15(6), pp.742-750, June, 2005.

[16] X. Zhang, et al, "Just-Noticeable Difference Estimation With Pixels in Images", Journal of Visual Communication and Image Representation, Vol 19(1), pp 30-41, 2008.
[17] C. H. Chou and C. W. Chen, "A Perceptually Optimized 3-D Subband Image Codec for Video Communication over Wireless Channels", IEEE Trans. Circuits Syst. Video Technol., vol. 6, No. 2, pp. 143-156, 1996.
[18] Z. Lu, et al, "Modeling Visual Attention's Modulatory Aftereffects on Visual Sensitivity and Quality Evaluation", *IEEE Trans. Image Processing*, Vol.14(11), pp.1928 – 1942, Nov. 2005.

[19] Z. Wang, et al, Foveation scalable video coding with automatic fixation selection. *IEEE Transactions on Image Processing*, 12:1703–1705, Feb 2003.
[20] W. Lin, et al, "Visual Distortion Gauge Based on Discrimination of Noticeable Contrast Changes", *IEEE Trans. Circuits and Systems for Video Technology*, 15(7):900- 909, July, 2005.
[21] J. Malo, et al, "Percetual feedback in multigrid motion estimation using an improved DCT quantization", IEEE Trans. Image Processing, 10(10):1411-1427, October, 2001.
[22] X. Yang, et al, "Just Noticeable Distortion Model and Its Applications in Video Coding", *Signal Processing: Image Communication*, 20(7):662-680, August 2005.
[23] Y. J. Chiu and T. Berger, "A Software-only Videocodec Using Pixelwise Conditional Differential Replenishment and Perceptual Enhancement", IEEE Trans. Circuits Syst. Video Technol., 9(3):438-450, April, 1999.
[24] R. J. Safranek, "A JPEG compliant encoder utilizing perceptually based quantization", SPIE Human Vision, Visual Proc., and Digital Display V, Vol. 2179, pp. 117-126, Feb. 1994.
[25] R. B.Wolfgang, et al, "PerceptualWatermarks for Digital Images and Video", Proc IEEE, 87(7):1108-1126, July 1999.

Weisi Lin graduated from Zhongshan Univ., China with B.Sc in Electronics and M.Sc in Digital Signal Processing in 1982 and 1985, respectively, and from King's College, London Univ., UK with Ph.D in Computer Vision in 1992. He taught and researched in Zhongshan Univ., Shantou Univ. (China), Bath Univ. (UK), Nat'l Univ. of Singapore, Institute of Microelectronics (Singapore), and Institute for Infocomm Research (Singapore). He served as the Lab Head, Visual Processing, and then the Acting Department Manager, Media Processing, in Institute for Infocomm Research. Currently, he is an Associate Professor in School of Computer Engineering, Nanyang Technological Univ., Singapore. His areas of expertise include image processing, perceptual modeling, video compression, multimedia communication, computer vision and embedded systems. He is a Chartered Engineer (UK), and serves a number of IEEE Technical Committees, as well as many international conferences. He is a member of the Board of Journal of Visual Communication and Image Representation.

In the areas related to visual perceptual modeling and processing, he holds 9 patents, wrote 2 book chapters, published 17 and 56 refereed papers in international journals and conferences, respectively, and made 10 contributions to international standardization. He has also been the project leader of 7 projects for perceptual visual processing, and maintained active, long-term working relationship with the companies which are keen in perception-based technology. He is the co-chair of the special sessions on perceptual processing in ICME06 and IMAP07. His recent major technical presentations include invited talk in VPQM06, the MAP keynote speech in ICCCN07, and tutorials in ISCAS08, PCM07 and ICME09, with different topics on visual perceptual processing.

Video Quality for Communications

Lina Karam, Arizona State University, USA
karam@asu.edu

Early communications applications were mainly focused on speech and audio communication. Increased computing performance and advances in communication networks and communication technologies were among the factors that led to more sophisticated communications applications supporting real-time image and video compression and transmission over various types of networks. This in turn led to an increasing demand for visual media communication in support of a variety of applications including, for example, wireless video, video conferencing, video surveillance, IPTV and, more recently, multi-view and 3D video applications.

Due to the limited available bandwidth, the video content has to be compressed before transmission, which typically degrades the perceived visual quality. Distortions in the visual quality are caused by various compression artifacts such as blockiness, blur, ringing, graininess, mosquito noise, ghosting, and jitter, to name a few. The quality of the compressed video can be further impaired during transmission over error-prone networks due to packet loss, packet drop, fading, and various other channel errors. These distortions can be very annoying to the consumers or can severely affect the performance of an application. In consumer-oriented applications, it is important to be able to assess the perceived video quality in order to provide the end user with the best possible experience subject to constraints on the available resources. In other applications, such as security, video surveillance, and medical imaging, assessing the video quality can help in ensuring a minimum level of quality that is needed for proper operation.

The ability to accurately assess the visual quality of transmitted video using objective metrics has recently gained a lot of interest. While subjective assessment experiments can be used for assessing the visual quality, these are complex and time-consuming. Furthermore, it is desirable to be able to perform real-time or frequent visual quality monitoring for video communication applications, which cannot be achieved with subjective assessments.

Objective video quality assessment metrics are designed to quantify the video quality automatically.

On one end of the spectrum, there are the full-reference objective quality metrics which require the availability of the original undistorted (reference) video. On the other end, there are the no-reference objective quality metrics which can quantify the video quality without knowledge of the original undistorted video but are much harder to design and can be less accurate as compared to the full-reference metrics. As an in-between tradeoff, reduced-reference objective quality metrics only require partial information about the original video.

Furthermore, objective quality metrics can require knowledge of the video compression scheme and/or the transmission environment. Some can be designed to only work with certain specific video codecs and/or transmission protocols. Alternatively, universal objective quality metrics can be developed to work without knowledge of the employed video compression and channel characteristics.

Objective quality metrics can be designed to measure specific visual impairments or they can be designed to assess the overall visual quality in the presence of various degradations. Equally important are metrics that can measure a contextual visual quality, or that can predict human perception as measured by the performance of a visual-based task.

The performance of objective quality metrics can be evaluated (initially before deployment) in terms of how well these metrics correlate with conducted subjective tests. Several performance evaluation metrics have been suggested by the Video Quality Expert Group (www.vqeg.org). Examples of performance evaluation metrics include the Pearson correlation coefficient, the Spearman rank-order correlation coefficient (SROCC), the mean absolute prediction error (MAE), the root mean square error (RMSE), and the outlier ratio (OR).

In the context of video communications, objective quality metrics can be used to monitor and control the quality of the transmitted video at the receiver end as well as at various locations within the network. This automatic quality monitoring would allow service providers to spot and resolve in a timely manner any arising problems and to maintain a competitive quality of service. The full-reference objective metrics are not suitable for real-time quality monitoring nor for low-delay video communications applications. In these cases, no-reference or reduced-reference metrics should be used. The full-reference metrics can possibly be used to perform delayed non real-time performance analysis. This can be done by collecting and storing locally, in real-time, sample partitions of the video during transmission at various network locations. Full-reference quality metrics can be computed using these collected samples as these samples can be retrieved by the transmitter (which has a copy of the original reference video) at a later time when a failure occurs for debugging or for collecting relevant statistics and performance data.

Reduced-reference and no-reference quality metrics can also be used for error detection and control. Conventional error control schemes are efficient at detecting bit errors but they do not provide information about the effects of a corrupted packet on the visual quality of the reconstructed video. Designing objective quality metrics that can measure these effects could result in effective visually-optimized error detection and correction schemes.

Reliable automatic assessment of video quality is still in its infancy and there is still a great need to develop objective quality metrics that can reliably assess the video quality in the presence of compression artifacts and channel errors. Research work is still needed for the development of reliable subjective and objective quality assessment methods, especially for high-definition video and 3D video. There is also a need to assess the quality of the overall user experience. This would require incorporating different modalities (such as auditory and affective in addition to visual) and possibly taking into account socio-economic factors and how these factors can affect the subjective quality ratings in the context of video communication applications.

Lina J. Karam received the B.E. degree in Computer and Communications Engineering from the American University of Beirut in 1989, and the M.S. and Ph.D. degrees in Electrical Engineering from the Georgia Institute of Technology in 1992 and 1995, respectively. She is currently a faculty member in the Electrical Engineering department at Arizona State University. Her research interests are in the areas of image and video processing, image and video coding, visual quality assessment, human perception, error-resilient source coding, medical imaging, and digital filtering.

Prof. Karam worked at Schlumberger Well Services on problems related to data modeling and visualization, and in the Signal Processing Department of AT&T Bell Labs on problems in video coding during 1992 and 1994, respectively. Prof. Karam is the recipient of an NSF CAREER Award. She is currently serving as an associate editor of the IEEE Transactions on Image Processing and as the Technical Program Chair of the 2009 IEEE International Conference on Image Processing. She served as the lead guest editor of the IEEE Journal on Selected Topics in Signal Processing, Special Issue on Visual Quality Assessment. She also serves on the editorial board of the Foundations and Trends in Signal Processing journal. She served as the Chair of the IEEE Communications and Signal Processing Chapters in Phoenix in 1997 and 1998. She also served as an associate editor of the IEEE Transactions on Image Processing from 1999 to 2003 and of the IEEE Signal Processing Letters from 2004 to 2006, and as a member of the IEEE Signal Processing Society's Conference Board from 2003 to 2005. She is an elected member of the IEEE Circuits and Systems Society's DSP Technical Committee and of the IEEE Signal Processing Society's IVMSP Technical Committee. She is a member of the Signal Processing, and Circuits and Systems societies of the IEEE.

Enhancing Video Quality: More MIPS or More BITS?

Gokce Dane and Khaled El-Maleh, Qualcomm, USA
gokced@qualcomm.com, kelmaleh@qualcomm.com

ABSTRACT

Digital video is one of the most attractive means of conveying information, and the quality of the video is one of the determining factors for the end user experience. In order to improve the visual user experience, one has to understand the quality expectation of users for the specific video application, and deliver upon good video quality or find ways to enhance it. The objective of this paper is to present various methods for video quality enhancement. We will review different methods available first from the encoder side and secondly from the decoder side.

CAUSES OF VIDEO ARTIFACTS

Video quality can be influenced from various factors starting from video capturing process, ending in display. The artifacts can be introduced at various stages of an end-to-end video system, including pre-processing, compression, transmission, post-processing and display. The video source can be noisy due to sensor characteristics of capturing device whether it is a single or multi-camera. Various compression artifacts such as blocking, blurring, ringing, color bleeding, flickering can be produced during encoding [1]. Video compression artifacts are varied by the available bit budget and the performance of the encoding algorithms. In addition to compression, quality can degrade severely due to possible losses during transmission. Tolerance to artifacts and acceptance of quality can vary depending on the device (whether it is a mobile phone or 3DTV) and the environment which the video is consumed, and lastly the expectation of user.

In order to improve the quality of video, the aforementioned artifacts have to be either reduced if not eliminated during encoding process or should be processed and masked after decoding. In the next sections we will address these two approaches.

MORE BITS: ENCODER SIDE

Video quality can be enhanced from the encoder side, simply by spending extra bits during compression. Spending more bits during compression gives the flexibility to encoder to provide better quality without increasing the complexity. On the other hand, encoder has the choice to enhance the quality by using more complex algorithms such as better motion estimation, better rate control and prediction schemes. However, even for very high-bit rate video compression applications such as HD video, where the bit-budget is very high, the given bit-budget has to be appropriately distributed among different scenes of a source video as well as on different regions of a frame. At the video sequence level, distributing the bits among different frames requires a good multi-pass rate control algorithm [2]. At the frame level, allocating bits to different regions of a video frame can be based on techniques such as ROI (region of interest) identification [3], and by using perceptual cues [4].

In the case of low bit rate video compression, one can send video content at full frame rate (30 fps) where the quality will be degraded by annoying blocking artifacts, due to the limited bit budget. In order to avoid compression artifacts, the frame rate may be reduced to 15 or 10 frames per second (fps) where each frame is encoded with more bits and therefore has better spatial video quality [5] as shown in the example in Figure 1. However in this case, a recovery mechanism utilizing frame rate up conversion is needed at the decoder to display the video at a higher frame rate [6], otherwise the video will suffer from motion jerkiness due to low frame rate. In another scenario, the video can be sent at full frame rate but at reduced resolution. This requires a good image down-sampling algorithm at the encoder.

In addition to temporal and spatial processing of video, color enhancement can be done at the encoder. Color enhancement is typically done by using enhanced chroma quantization in regions such as bright red areas where users pay more attention to, or between areas of strong chrominance difference to avoid color bleeding. This will also require more bits for coding color information.

(i)

(ii)

Figure 1 Comparison of 2 frames (i) QCIF 10 fps encoded at 48 kbps (ii) QCIF, 5 fps encoded at 48 kbps.

To enable the compressed bit-stream to resist channel errors which might occur during transmission error resiliency techniques can be applied during encoding, so that the impact of errors on reconstructed video quality is minimal. For error resiliency, encoder has to add additional redundancy to the bit-stream and this can take up to 20% of overall complexity. Many tools such as reference frame selection, flexible macroblock ordering, intra-block refresh to name a few exist in H.264 standard [7].

In summary, encoders can perform analysis and determine how to spend the bits smartly for the best quality. They have the option to increase quality without complexity. Furthermore, for a given bit budget, they can increase the quality further by using more complex algorithms. Overall, the approach of enhancing the quality at the encoder is *good for all video decoders* if the bit-stream produced by encoder is compliant to the standard which the decoder can support, and no standard-noncompliant side information is sent.

MORE MIPS: DECODER SIDE

Another perspective is to enhance the quality at the decoder side, simply by spending more MIPS, i.e. additional resources during decoding, or after decoding as a post-processing stage.

In order to eliminate compression artifacts, various enhancement techniques such as *deblocking, de-ringing, color bleeding, de-blurring/sharpening* can be applied at the decoder side. Since the original reference frame is not available, the decoder has to first identify the locations of artifacts reliably and perform high-quality correction. In case of quality degradations due to losses during transmission, error concealment techniques that utilize motion vectors could be applied [8].

If the encoder sends a lower resolution image, image up-sampling techniques can be applied at the decoder to increase the resolution. The quality of up-sampled image depends on the particular interpolation technique that decoder uses. Interpolation could be a simple bilinear method, or a complex super-resolution based method [9]. One example of the quality enhancement effect of up-sampling at the decoder can be seen in Figure 2.

(i)

(ii)

Figure 2 Comparison of two frames (i) QVGA, 15 fps encoded at 48 kbps (ii) QCIF, 15 fps encoded at 48 kbps and up-sampled to QVGA at the decoder.

The visibility of compression artifacts vary based on the resolution of the content and the distance of the observer to the display [10]. For example, users are more willing to accept more distortions in small screens, but the tolerance is much lower for higher resolutions and bigger display sizes. Similarly 24 fps video might be sufficient for mobile, but high-end TVs require higher frame rates for better smooth motion perception, and motion blur reduction [11]. Whether it is mobile or high-end TV, frame rate up conversion techniques can be applied to enhance the temporal video quality.

Besides quality of motion and picture resolution, color vividness is another factor that determines the user's acceptance of the content. Color enhancement requirements are different for mobile and large-area displays [12]. Color could be enhanced on specific regions considering human visual system (HVS) by using existing color gamut, or the whole color gamut of a display could be extended.

Even if the video content might be free from artifacts, there are still various techniques that could be applied by the display system such as contrast enhancement, high dynamic range adjustment, or backlight adjustment to name a few, which will increase the user experience of the digital video content.

In summary, improving the video quality at the decoder and display is *good for all type of video encoders*, but this approach can increase the complexity of the decoder immensely depending on the type of algorithm applied.

CONCLUDING REMARKS

In this paper, we addressed the problem of video quality enhancement from two perspectives, and listed variety of methods available in the literature and applied in industry. These techniques could work independently, making either encoder or decoder computationally loaded, or can work jointly. Finally, we pose two important questions:

- For a given bit budget, can the video quality of a higher bit rate be achieved by just using post-processing?
- What kind of post-processing is required to achieve such higher quality?

For certain applications, and certain content (such as the one application shown in Figure 2), obtaining the quality of higher bit-rate can be possible. But more research and investigation is necessary to find the limits of post-processing algorithms.

REFERENCES

[1] M. Yuen, H. R. Wu, "A survey of Hybrid MC/DPCM/DCT Video Coding distortions," *EURASIP Journal on Signal Processing*, vol. 70, 1998.

[2] P. H. Westerink, R. Rajagopalan, C. A. Gonzales, "Two-pass MPEG-2 variable-bit-rate encoding", *IBM Journal of Research and Development*, Vol. 43, Number 4, 1999.

[3] H. Wang, Y. J. Liang, K. El-Maleh, "Real-time region-of-interest video coding using content-adaptive background skipping with dynamic bit allocation", *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2006.

[4] I. Ahmad, J. Luo, "On using game theory to optimize the rate control in video coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no.2, 2006.

[5] H. Song, C.-C. Kuo, "Rate control for low-bit-rate video via variable-encoding frame rates", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no.4, 2001.

[6] G. Dane, T. Nguyen, "Optimal temporal interpolation filter for motion-compensated frame rate up conversion", *IEEE Transactions on Image Processing*, vol. 15, no.14, 2006.

[7] T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, 2003.

[8] G. Dane, Y. Ye, Y.-C. Lee, "Low-complexity temporal error concealment by motion vector processing for mobile video applications", *Proceedings IEEE International Conference on Image Processing*, 2008.

[9] W. T. Freeman, T. R. Jones, E. Pasztor, "Example-based super-resolution", *IEEE Computer Graphics and Applications*, March/April, 2002.

[10] S. H. Bae, T. Pappas, B.-H. Juang, "Subjective evaluation of spatial resolution and quantization noise tradeoffs", *IEEE Transactions on Image Processing*, vol. 18, no. 3, 2009.

[11] J. H. Souk, J. Lee, "Recent picture quality enhancement technology based on human visual perception in LDC TVs", *Journal of Display Technology*, vol. 3, no. 4, 2007.

[12] L. D. Silverstein, "Color display technology: from pixels to perception", IS&T Reporter "The window on imaging" vol. 21, no. 1, 2006.

Gökçe Dane (S'01–M'05) received the B.S. degree in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 1999, the M.S. degree in electrical engineering from the University of California, Riverside, in 2001, and the Ph.D. degree in electrical and computer engineering from the University of California-San Diego, La Jolla, in 2005. She was with HP Labs, Palo Alto, CA, and Qualcomm, Inc., San Diego, during the summers of 2000 and 2004, respectively. She worked in Thomson Corporate Research in Princeton between 2005 and 2006 on a professional H.264 based high definition video encoding system for use in HD-DVD and Blu-Ray disc products. She is currently with Qualcomm Inc, QCT Display Systems Group in San Diego. Her research interests are in the general areas of signal, image, and video processing, compression, and multimedia communications. Her current work includes developing various pre/post/display processing algorithms and temporal quality enhancement techniques for mobile platforms.

Khaled El-Maleh (S'94–M'00) received PhD and M. Eng. From McGill University, Canada, and B Sc. in EE and B. Sc. in Applied Math from King Fahd University (KFUPM), Saudi Arabia. He has been with Qualcomm, San Diego since 2000. Khaled areas of expertise and interests include: design, implementation and quality evaluation of mobile multimedia systems with focus on speech and video compression\ processing algorithms. Khaled was a key member of the team who designed the first generation of mobile camcorder and video telephony solutions for both UMTS and CDMA networks. He was the video quality lead for Qualcomm chip unit. He has keen interest to build stronger bridges of collaboration between academic institutes and industry with focus on innovation. In addition to his technical publications, he filed more than 200 US and international patents covering video\image\ speech processing and compression technologies.

Quality-of-Service Support for Next Generation Satellite Multimedia Networks: Opportunities and Challenges

Hongfei Du, Jiangchuan Liu, Jie Liang, Simon Fraser University, Canada

hongfei.du@ieee.org, jcliu@sfu.ca, jiel@sfu.ca

Introduction

Satellite networks are particularly appealing for reducing the "digital divide" when the remote users need to be connected to the terrestrial backbone infrastructure or when interconnecting isolated regional networks. However, satellite-based systems are faced with severe propagation impairments, long latency, constrained bandwidth and power. Such issues become more challenging when delivering triple-play services with broadcast/multicast (BC/MC) access. Moreover, multimedia applications have been evolving and play an overwhelming role in today's Internet traffic. Such applications pose stringent requirements in terms of quality-of-service (QoS), e.g., delay, bandwidth, and jitter. The satellite-delivered multimedia offer promising opportunities as well as crucial hurdles for heterogeneous content provisioning in spectrum-efficient and economic-viable manner. In this article, we review current worldwide advances in satellite systems for multimedia broadcast/multicast. We found that a key component is employed in many of these systems for complementary indoor/in-building coverage, i.e., the gap-filler, also known as complimentary ground component (CGC) or intermediate module repeater (IMR). In particular, we address the quality-of-service (QoS) aspects for the integrated terrestrial-satellite multimedia system and relevant optimization techniques, thereby igniting the future research directions in the field.

The State of the Art

Research and development in satellite-delivered multimedia communication systems can date back to 1990s, where Asynchronous Transfer Mode (ATM) based satellite system was developed in Europe RACE II CATALYST project for the provisioning of Broadband ISDN (B-ISDN) services, supporting a variety of multimedia services and different LAN architecture [1]. Very Small Aperture Terminal (VSAT) was one of the most successful satellite communication systems that attempts to support multimedia contents provisioning; it defines a class of very small aperture (0.75-2.4m) to make the multimedia application available at homes and offices. Both approaches demonstrated their capabilities of supporting multimedia communications via satellite platform, yet such attempts are less than successful in that the available satellite bandwidth is much less than that offered by the terrestrial networks and therefore is not sufficient for the rapidly growing demand of multimedia applications. As such, it is essential to optimize the bandwidth efficiency while taking into account different aspects from the protocol design to system architecture. During the past years, various initiatives in Europe, such as Multimedia Broadcast Multicast Services (MBMS) [2], satellite UMTS (S-UMTS), satellite digital multimedia broadcasting (SDMB) [3], digital video broadcasting via satellite handheld (DVB-SH), and European Satellite Digital Radio (ESDR), have made impressive progresses in developing feasible, scalable and flexible protocols and systems, with several attempts on providing the high bandwidth efficiency as well as QoS guarantee for satellite-delivered multimedia services.

Amongst those efforts, the SDMB system has been a key subject of EU funded projects. The early development of the concept of SDMB is emerged from research efforts in the S-UMTS. This subject is subsequently extensively studied to cover aspects of the system definition, business opportunities, implementation, performance evaluation and validation, via several EU R&D projects including SATIN, MODIS [4] and MAESTRO [5]. Representative systems that are most likely to be commercially deployed include EUROPA-MAX, ONDAS, Eutelsat/SES SOLARIS, and MAESTRO end product "Unlimited Mobile TV". By using three HEO satellites with limited number of gap-fillers, the ONDAS system is able to provide high-quality radio and video contents to users across Europe. Operated in the S-Band (2170-2200MHz), the SOLARIS is expected to supply 60 channels using 6 spot beams, while the number of channels can be tripled when the terrestrial gap-fillers is added. The "Unlimited Mobile TV" employs the DVB-SH technology, which is currently being standardized by the DVB Forum. The solution would be capable of delivering up to 45 channels at 256 kbps throughout Europe. Other project like SECOM/ABATE, adopts the high frequency Ka and EHF bands, which enjoy large spectrum resources and reduced antenna dimension in both terminal and spacecraft. It is expected that the first commercial SDMB services would be available in the next couple of years. In the USA, similar systems have been developed, e.g., XM Radio, Sirius Satellite Radio, and ICO's Mobile Interactive Media (MIM). Notable solutions in other parts of the world include MBSAT (Mobile Broadcasting Satellite) in Korea, MobaHO (Mobile Broadcasting Service) in Japan, and CMMB (China Multimedia Mobile Broadcasting) in China.

Opportunities and Challenges

The advantages of satellite based transmission lie in its broadcast nature and extensive coverage without involving additional deployment cost. Therefore, the satellite appears as the single fundamental component that can provide cost-economical ubiquitous content access to anywhere in the world. However, challenges remain in the diverse aspects:

- Satellite link inevitably involves tough delay and loss model, i.e., long delay and more transmission errors.
- Limited bandwidth and power.
- Services are intended for multiple users with different location and channel statuses.
- Return link diversity: either via satellite component or via terrestrial network.

These characteristics render fast closed-loop power control (CLPC) and effective channel state information (CSI) and end-to-end (ETE) measures unavailable. Among the striking features of integrated terrestrial-satellite infrastructure, is the fact that no single entity is responsible for the end-to-end performance assurance. Indeed, the performance of satellite-delivered multimedia does not always measure up to its contracted targets, and the final delivered QoS remains subpar. Providing QoS assurance for satellite-delivered multimedia is challenging in that QoS support in satellite multimedia network not only concerns the network-centric QoS performance, such as throughput, delay and jitter, but also the application-centric and user-centric metrics, in accordance with the system bandwidth/power constraints and the user reception conditions. Such a challenge is of special interest for multi-session broadcast/multicast via long-latency geostationary bent-pipe satellite, where little remarkable effort can be noted so far.

Broadcast/Multicast Access

Another concern for satellite-delivered multimedia is its native BC/MC support for multi-session heterogeneous service provisioning. It increases dramatically the transmission capacity, yet poses challenges on appropriate settings of transmit bandwidth/power, as a single session at the gateway usually corresponds to multiple receivers in a spot-beam area, each feature diverse and fast-varying capacities and reception conditions. To cope with the highly vibrating satellite fading channels, the CSI information and ETE metrics from respective BC/MC members should be taken into account in the protocol design. Let us assume a feedback report is perfectly generated at the receiver and is reliably fed back to the gateway, reporting the current CSI and ETE conditions. Upon receiving the feedback information from each BC/MC member in the intended BC/MC group in a dynamic and periodical manner, the gateway subsequently derives the overall reception level for each BC/MC session associated with the entire BC/MC group. There can be various ways for measuring the overall BC/MC group reception level. A simple approach may measure the worst-case or best-case reception conditions, or some melding in between. Nonetheless, it may not perform optimum for the reason that the same transmit rate/power at the gateway will not scale well for multiple receivers with diverse rate/power expectations. Another viable approach would measure the overall BC/MC group performance as the instantaneous percentage of members given good reception condition.

Return Link Diversity

The presence of terrestrial return channel facilitates interactive activities for satellite-delivered multimedia. It is worth noting that there exist major discrepancies on whether a return link via terrestrial or via satellite is used.

- Via terrestrial: for the BC/MC members with an accessible terrestrial return link, each gap-filler performs the initial gathering of the channel status for all BC/MC members in the group and reports it altogether to the gateway.

- Via satellite: for the BC/MC members with direct access signal, the satellite link will be the only way for the gateway interworking with the remote BC/MC members.

To effectively manage the radio resources and maximize the channel capacity, return link adaptation upon differentiated return link mode should be considered. Besides, gap-fillers play a key role in maintaining the reliability and scalability of the overall system performance. Apart from its simple forwarding functions, it is desired to conduct the measurements and assessments on reception status of all the BC/MC members in its cell, and then report the overall status to the gateway.

Conclusions

This article discusses the state of the art on multimedia QoS support for future satellite systems. Satellite-delivered multimedia is promising yet still in its early stage. We envision the key design issues pertinent to the end-to-end multi-session QoS performance assurance in long-latency bent-pipe broadcast/multicast satellite systems. We argue that efficient and adaptive protocol design in such a challenging scenario is desired to incorporate multiple essential factors, e.g., the reception conditions, and return link diversities. Considerable research and development work is needed in assuring comprehensive QoS demands of satellite-delivered multimedia applications.

References

[1] Sun, Z. Hadjitheodosiou, M.H. Komisarczuk, P. Coakley, F.P. “Satellite connections for multimedia communications,” IEE Multimedia Communication Systems, Mar 1994, pp. 5/1-5/6.

[2] 3GPP TS 23.246 v6.8.0, “Multimedia Broadcast/Multicast Service; Architecture and functional description, (Release 6),” October, 2005.

[3] N. Chuberre, et al, “Relative Positioning of the European Satellite Digital Multimedia Broadcast (SDMB) Among Candidate Mobile Broadcast Solutions,” *IST Mobile & Wireless Commun. Summit*, Germany, 19-23 June, 2005.

[4] EU FP5 IST MoDiS, http://www.ist-modis.org.

[5] EU FP6 IST MEASTRO, http://ist-maestro.dyndns.org.

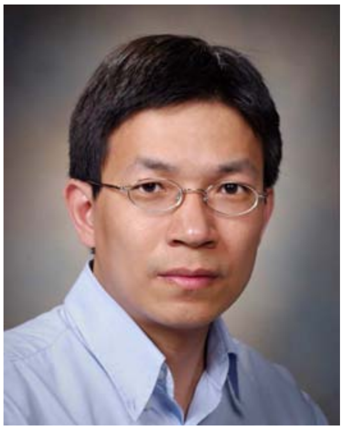
Hongfei Du received the B.Eng degree in electronic information engineering from the Department of Electronic Engineering, Beijing University of Aeronautics & Astronautics, Beijing, China, in 2003. He received the M.Sc, M.Phil, and Ph.D degrees in mobile and satellite communications, from the Center for Communications Systems Research (CCSR), School of Electronic and Physical Science, University of Surrey, Guildford, Surrey, United Kingdom, in 2004, 2005 and 2007, respectively. From Aug 2007, he joint the BROADWAYS group at CREATE-NET international research institute, Trento, Italy, as research consultant, coordinating and conducting EU research project on middleware design for mobile multimedia broadcasting.

From Aug. 2008, he is with Simon Fraser University, as a postdoctoral research fellow on efficient video compression/transmission over wireless mesh networks. Hongfei Du has been involved in numerous research projects in the area of mobile and satellite communications systems. His main research interests include multimedia broadcasting, scalable video coding, radio resource management, packet scheduling, quality of service support and cross-layer design.

Jiangchuan Liu (S'01-M'03-SM'08) received the BEng degree (cum laude) from Tsinghua University, Beijing, China, in 1999, and the PhD degree from The Hong Kong University of Science and Technology in 2003, both in computer science. He was a recipient of Microsoft Research Fellowship (2000), a recipient of Hong Kong Young Scientist Award (2003), and a co-inventor of one European patent and two US patents. He co-authored the Best Student Paper of IWQoS'08.

He is currently an Assistant Professor in the School of Computing Science, Simon Fraser University, British Columbia, Canada, and was an Assistant Professor in the Department of Computer Science and Engineering at The Chinese University of Hong Kong from 2003 to 2004.

His research interests include multimedia systems and networks, wireless ad hoc and sensor networks, and peer-to-peer and overlay networks. He is an Associate Editor of IEEE Transactions on Multimedia, and an editor of IEEE Communications Surveys and Tutorials. He is a Senior Member of IEEE and a member of Sigma Xi.



Jie Liang (S'99-M'04) received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, in 1992 and 1995, the M.E. degree from National University of Singapore (NUS), in 1998, and the Ph.D. degree from the Johns Hopkins University, Baltimore, MD, in 2003, respectively. Since May 2004, he has been an Assistant Professor at the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. From 2003 to 2004, he was with the Video Codec Group of Microsoft Digital Media Division, Redmond,WA. His research interests include image and video coding, multiview video coding, multirate signal processing (wavelets and filter banks), multimedia communications, and information theory. He is a Professional Engineer in British Columbia.

Editor's Selected Paper Recommendation

B. Han, L. Ji, S. Lee, B. Bhattacharjee, and R. R. Miller, "All bits are not equal – A study of IEEE 802.11 communication bit errors," *IEEE Infocom'09*, April 19-25, 2009, Rio de Janeiro, Brazil.

It is well It is well known and recently analyzed as in articles [1] and [2] that not all packets in video applications are equal. Rather packets, dependent on the used video coding scheme, generally have different importance to and impact on the viewer's quality of experience (QoE). Technologies such as forward error correction (FEC) and retransmission-based loss recovery can be employed to defeat packet loss and enhance the QoE. Even multiple erroneous receptions of a given frame can be combined together to recover the original frame/packet without further retransmission [3].

Another similar question is "are all bits of a packet equal?" or "will all bits of a packet experience identical loss probability?". The Infocom'09 paper (authored by B. Han, L. Ji, S. Lee, B. Bhattacharjee, and R. R. Miller, from University of Maryland and AT&T Labs – Research, with the title "All bits are not equal – A study of IEEE 802.11 communication bit errors"), as the first work, answers NO to this question. This paper presents experimental results obtained from a study focusing on wireless local area networks (WLANs) transmission bit errors and discovers some unique transmission bit error characteristics in WLANs.

In this work, the authors studied WLAN transmission errors on the "sub-frame" level. They have conducted extensive experiments on several IEEE 802.11 wireless LAN testbeds. The measurement results have identified that in addition to bit error distribution induced by channel condition, other bit error probability patterns also exist across different communication environments and different hardware platforms. Three patterns for bit error probabilities with respect to bit position in a frame have been identified, namely the slope-line pattern, the saw-line pattern, and the finger pattern, which may not be caused by channel fading. First, there is a linear relationship between the frequency of bit errors and the bit position in the frame. A bit near the end of a frame is more likely to be corrupted than a bit near the beginning of the frame. It's referred to the slope-line pattern in this paper. The saw-line pattern is the fine zig-zag line that goes across the full length of the frame with the saw-tooth peak-to-peak period about the same as the number of bits each OFDM symbol carries at a given transmission bit rate. The finger pattern, i.e., the larger peaks, starts to appear after certain bit position and repeats at a fairly regular interval. Please see figure 1, abstracted from this paper.

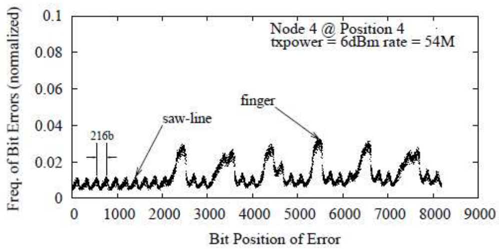


Figure 1. Normalized bit error frequency for a given receiver node with data rate 54 Mbps (The average RSSIs of correct packets, truncated packets and packets with bit errors are 36, 21 and 22, respectively).[4]

While it is challenging to figure out the exact causes of these patterns without access to detailed WLANs hardware designs, they provided some possible reasons: clock drift and changes of channel condition for slope-line pattern, the frequency selectivity characteristic of wireless channel for the saw-line pattern, and the inter-play between the transmitter's power control loop and the receiver's gain control loop for the finger pattern.

Such repeatable and predictable patterns discovered in this paper can be exploited for designing more efficient sub-frame or sub-packet level mechanisms such as frame combining to improve the performance of multimedia applications, such as video streaming over WLANs. Another interesting question to ask is "would those patterns appear in other wireless communication technologies such as ZigBee?"

References

[1]. J. Greengrass, J. Evans, and A. C. Begen, "Not all packets are equal, part 1: Streaming video coding and SLA requirements," IEEE Internet Computing, vol. 13, no. 1, Jan.-Feb., 2009, pp. 70-75.

[2]. J. Greengrass, J. Evans, and A. C. Begen, "Not all packets are equal, part 2: The impact of network packet loss on video quality, " IEEE Internet Computing, vol. 13, no. 2, Mar.-Apr., 2009, pp. 74-82.

[3]. A. Miu, H. Balakrishnan, and C. E. Koksal. "Improving Loss Resilience with Multi-Radio Diversity in Wireless Networks," In Proceedings of the ACM MOBICOM 2005, pages 16-30, August-September 2005.

[4]. B. Han, L. Ji, S. Lee, B. Bhattacharjee, and R. R. Miller, "All bits are not equal – A study of IEEE 802.11 communication bit errors," IEEE Infocom'09, April 19-25, 2009, Rio de Janeiro, Brazil.

** The Column Editor recommending this paper is Dr. Chonggang Wang.*

Focused Technology Advances Series

Distributed Signal Processing in Wireless Embedded Networks

Antonios Argyriou, Phillips Research, Netherlands
anargyr@ieee.org

The vision of ambient intelligence consists of a multitude of electronic devices and sensors that are seamlessly embedded into people's daily life. Currently, the most promising applications for this environment include home entertainment, healthcare, monitoring, automation, while it is not clear yet what innovative applications will surface in the future. In this article, we touch one aspect related to one discipline that is expected to be crucial for the realization of this vision. More specifically, we discuss signal processing in distributed and embedded wireless networks. There are several reasons for studying more closely aspects of the signal processing application and its implementation in embedded wireless devices.

One of the basic characteristics of this environment is the plethora of the physical sensing devices and the heterogeneity of the information sources. Consider for example sensor nodes that monitor the temperature, presence, light, or even information coming from the Internet etc. For this information to be useful, different signal processing algorithms need to be employed. Sensor signal processing researches the use of statistical methods for extracting useful information for the detection, characterization and recognition of variables in the noisy environment. Namely three major areas of statistical signal processing are usually employed in sensor networks: statistical detection and estimation; adaptive-signal processing; and signal classification. The interesting aspect that we have to pay attention is that the complexity of the aforementioned processing algorithms sets different requirements on the underlying hardware. For example several signal processing applications require a certain sampling rate in order to produce a meaningful output. In cases like this, tradeoffs between the use of specialized hardware (ASIC), a digital signal processor (DSP), or even general purpose micro-controllers are really important. Furthermore, signal processing applications impose different requirements on the underlying wireless communication technology since data may have to be delivered in a timely fashion. Finally, one of the most important concerns for embedded applications is low power consumption and battery life. Squeezing even the last possible energy bit is important. Therefore, we believe that holistic system design methodologies are needed.

Another important characteristic of this environment of sensors and devices is that it is networked, distributed, and self-organized. This means that sensing devices usually have to be able to be deployed independently and use wireless communication for exchanging information between each other. Therefore, multi-hop wireless connectivity is needed in order to enable communication not only with the infrastructure-based network but also between sensing devices themselves. At the network layer, the wealth of research in the area of routing protocols wireless ad hoc networks can find its way into practical applications. However, one aspect that we believe that needs further investigation is the scalability of existing protocols in large scale deployments. At the medium access control (MAC) layer low power operation is of paramount importance. The MAC protocol has to be adaptive to the signal processing application because of the highly irregular and application-specific data traffic patterns. If we look the problems at the highest layer, applications in this environment must be able to use intelligently the information that is collected by the multiple sensors by cooperating. In this case we move from the problems related to distributed signal processing to the concept of distributed reasoning. Therefore, information/data fusion is very important since there are different types of sensors and applications.

One can easily see that several cross-domain problems still have to be investigated. The context in which the aforementioned issues have to be addressed is system-wide while application-specific metrics need to have more central role. However, the success of specific wireless embedded signal processing systems is likely that it will be determined by the useful applications that can be supported.

Antonios Argyriou received the Diploma in electrical and computer engineering from Democritus University of Thrace, Greece, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering as a Fulbright scholar from the Georgia Institute of Technology, Atlanta, USA, in 2003 and 2005, respectively. Currently, he is a Senior Research Scientist at Philips Research, Eindhoven, the Netherlands. From 2004 until 2005, he was also a research engineer at Soft.Networks, Atlanta, GA, and was involved in the design of MPEG-2 and H.264 video codecs. His current research interests are in the areas of wireless communications and networking. He has published more than 30 papers in international journals and conferences. He is a member of IEEE.

IG Corner: Seamless Mobility in Wireless Networks

Jiang (Linda) Xie, University of North Carolina-Charlotte, USA
Xiaoyuan Gu, Massachusetts Institute of Technology, USA
jxie1@uncc.edu, gu@sloan.mit.edu

The ever-increasing demands for mobile services have called for seamless mobility in wireless networks. Seamless mobility offers transparent mobile services to users while they are moving in wireless networks. Seamless mobility support is an old research topic, but an important and indispensable component of wireless networks. With the advances of wireless technologies and the continuing growing demand of mobile services, seamless mobility support is facing new challenges. Traditional approaches for providing seamless mobility are not sufficient to tackle all the challenges manifested today.

Challenge 1 --- Emerging Wireless Technologies: several new wireless technologies have been introduced in recent years, such as WiMAX, LTE, wireless mesh, and cognitive radio technologies. These emerging wireless technologies bring new challenges for seamless mobility support and hence push the seamless mobility issue in wireless networking environments to new domains including mobility in multihop wireless networks, mobility in infrastructureless networks, and frequency-agile mobility.

Challenge 2 --- Coexistence of Heterogeneous Wireless Technologies: the evolution of many wireless access technologies has enabled the realization of a wide range of wireless networks, such as traditional and next-generation cellular networks, wireless local area networks (WLANs), wireless personal area networks, wireless body area networks, wireless sensor networks, vehicular ad hoc networks, and satellite networks. While no single wireless technology is predominant today, no single wireless technology will prevail in the foreseeable future, due to the fact that different wireless technologies were designed to address different coverage, mobility, and data rate requirements. The coexistence of heterogeneous wireless technologies is becoming an unavoidable reality as many more new mobile devices become equipped with multiple and heterogeneous wireless interfaces. This brings a lot of opportunities and challenges for seamless mobility in heterogeneous wireless environments.

Challenge 3 --- New Multimedia Services and Applications: new types of multimedia services are becoming more popular in recent years, such as voice over IP (VoIP), multimedia messaging, video on demand, IPTV, online music/video downloading, gaming, and location-based services. In addition, new applications using wireless technologies are also emerging including context-aware networking, social networking, e-commerce, tele-medicine/e-health, etc. As technologies advance in providing improved low-latency and high-capacity mobile broadband environments, consumers would expect from their mobile devices to receive all the multimedia services they can receive from landline access. This adds a new dimension of complexity for seamless mobility support in providing guaranteed quality of service for these new multimedia services and applications in mobile environments.

Challenge 4 --- Changing Mobility Characteristics: with the introduction of new mobile services and devices, online users have changed their behaviors. For example, users carrying lightweight mobile devices (e.g., iPhone) and requesting VoIP services display more mobile behaviors and have longer session durations, compared to traditional WLAN laptop users. This changing of mobility characteristics for emerging mobile services affects the prediction accuracy of user mobility behaviors, and hence affects mobile networking protocol design and seamless mobility support.

Issues to Address: On the one hand, it is of critical importance that seamless, low latency, and transparent services be provided to mobile users, via potentially multiple heterogeneous wireless technologies or opportunistic spectrum access involved during the course of movement. Therefore, ***mobility management*** issues need to be addressed adequately. Mobility management includes a set of management mechanisms that enable the network to maintain connections as a mobile terminal is moving into a new service area and to locate a roaming terminal for packet delivery. Thus, with the support of mobility management, mobile terminals can freely roam with uninterrupted services, enjoying ubiquitous wireless access. Issues related to mobility management for seamless mobility support include:

- Vertical handoff and seamless integration of heterogeneous networks;
- Location tracking, positioning, and address management;
- Environment cognizance, spectrum-awareness, and location-awareness during mobility;
- Frequency agility associated with mobility and spectrum handoff;
- Security and privacy issues during mobility, including authentication, key management, trust models, mobility-related signaling message protection, etc.

On the other hand, host and network mobility also affects the performance of networking protocols significantly, which in turn makes ***mobility adaptability*** an important design issue. This asks the design of wireless architectures, protocols, spectrum management, and mobility management mechanisms to be revisited. Issues related to mobility adaptability include:

- Mobility modeling for emerging applications;
- Mobility-aware adaptive and resilient MAC, routing, power control, and spectrum management protocols;
- Opportunistic interconnections of heterogeneous wireless networks;
- Quality-of-Service (QoS) adaptation during mobility;

While these issues have received some attention in recent years and various standard bodies have put efforts in addressing these issues, including the mobility management schemes for infrastructure networks that are under the development of the Internet Engineering Task Force (IETF) and the IEEE 802.21 (Media Independent Handover) to facilitate seamless handovers in infrastructure wireless networks (e.g., IEEE 802.11, IEEE 802.16, and 3GPP), most work to date has not addressed specifically the new challenges mentioned above. More recently, the U.S. National Science Foundation (NSF) sponsored an academic workshop on wireless mobility to discuss future research opportunities in terms of adaptive protocols and systems under mobility.

In summary, seamless mobility in wireless networks is a challenging and important issue. Providing seamless mobility to emerging multimedia services with new wireless technologies incurs even more challenges and may require a different mindset from traditional approaches. With the support of seamless mobility, the applicability of many wireless applications can be significantly enlarged to mobile environments.

For more information:

[1] US National Science Foundation (NSF) workshop: *Mobility in Wireless Networks: Challenges and Opportunities*, July 2007, http://netlab.cs.ucla.edu/mwnet/usemod10/wiki.cgi?Main

[2] J. Xie and X. Wang, "A survey of mobility management in hybrid wireless mesh networks," *IEEE Network*, vol. 22, no. 6, pp. 34-40, November/December 2009. (Ranked 18th of Top 100 Documents Accessed in IEEE Xplore, 6th of Ten Most Popular Articles published in IEEE Communications Society periodicals viewed online, in January 2009)

Jiang (Linda) Xie serves as the Vice-Chair of the Interest Group on Autonomic Communications (ACIG), IEEE Communications Society Technical Committee on Multimedia Communications (MMTC). She received her PhD degree from Georgia Institute of Technology in 2004. Currently, she is an assistant professor in the Department of Electrical and Computer Engineering, University of North Carolina at Charlotte. Her research interests lie in the field of wireless networking with emphasis on architecture and protocol design for mobility management, radio resource management, and Quality-of-Service management for the next-generation wireless Internet. She is on the editorial board of

Computer Networks journal (Elsevier), *Journal of Network and Computer Applications* (Elsevier), and *Journal of Communications* (Academy Publisher). She has been on the technical program committee of different IEEE conferences and chaired some of their sessions. Currently, she is a Symposium Co-Chair of Wireless Networking Symposium of Globecom 2009 and Globecom 2010.

Xiaoyuan Gu is a System Design & Management Fellow at the Massachusetts Institute of Technology where he is currently working on an Engineering and Management Degree (SM). Previously, he was a research staff member at the Institute of Operating Systems and Computer Networks in Technical University of Braunschweig, Germany. He received his MSc from the International University Bruchsal, Germany in 2001, and is finishing up his PhD in Computer Science from TU Braunschweig, Germany. Between 2001 and 2003, he was with the Panasonic Multimedia Communications European Lab as a research engineer. His research interests include autonomic communications, network architecture, engineering management and technology strategy. He co-chairs the IEEE Interest Group on Autonomic Communications and has served on the Technical Program Committees of a number of major conferences such as IEEE Globecom, ICC, and CCNC.

MMTC COMMUNICATIONS & EVENTS

Call for Papers of Selected Journal Special Issues

Journal of Communications
Special Issue on Network Coding and Applications

Guest Editors: Antonios Argyriou, Joerg Widmer, Bob Li
Paper Submission deadline: **April 30**, 2009
Target Publishing Issue: 4th Quarter, 2009
CfP Weblink: http://www.academypublisher.com/jcm/si/jcmsi_nca.html

IEEE Network Magazine
Special Issue on Digital Home Services

Guest Editors: Antonio Sánchez-Esguevillas, Belén Carro-Martínez, Bin Wei, Wen Feng
Paper Submission deadline: **May 1**, 2009
Target Publishing Issue: Nov./Dec. 2009

International Journal of Digital Multimedia Broadcasting
Special Issue on Network-Aware Peer-to-Peer (P2P) and Internet Video

Guest Editors: Ken Kerpez, Yuanqiu Luo, Stan Moyer, John Buford, Dave Marples
Paper Submission deadline: **June 1**, 2009
Target Publishing Issue: December 1, 2009
CfP Weblink: http://www.hindawi.com/journals/ijdmb/si/p2p.html

Journal of Communications
Special Issue on Multimedia Computing and Communications

Guest Editors: Fan Zhai, Homer Chen, Thomas Stockhammer, Touradj Ebrahimi
Paper Submission deadline: **September 1**, 2009
Target Publishing Issue: 2nd Quarter, 2010
CfP Weblink: http://www.academypublisher.com/jcm/si/jcmsi_mcc.html

International Journal of Digital Multimedia Broadcasting
Special Issue on Video Analysis, Abstraction, and Retrieval: Techniques and Applications

Guest Editors: Jungong Han, Ling Shao, Peter H.N. de With, Ling Guan
Paper Submission deadline: **September 1**, 2009
Target Publishing Issue: March 1, 2010
Cfp Weblink: http://www.hindawi.com/journals/ijdmb/si/varta.html

Call for Papers of Selected Conferences

MINES 2009	
Website:	http://liss.whu.edu.cn/mines2009/
Dates:	Nov. 18 - 20, 2009
Location:	Wuhan, China
Submission Due:	**May 1, 2009**

NPSEC 2009	
Website:	http://www.gtisc.gatech.edu/npsec09/
Dates:	Oct. 13, 2009
Location:	Princeton, USA
Submission Due:	**May 29, 2009**

CCNC 2010	
Website:	http://www.ieee-ccnc.org/2010/
Dates:	Jan. 9, 2010
Location:	Las Vegas, USA
Submission Due:	**June 1, 2009**

Next Issue Partial Content Preview