DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

36-545 MIT, CAMBRIDGE, MASSACHUSETTS 02139-4307

William F. Schreiber
Professor of Electrical Engineering,
Emeritus

DOCKET FILE COPY ORIGINAL

RECEIVED
JUL 11 1996
FCC MAIL ROOM

10 July 1996

Secretary, Federal Communications Commission
1919 M St. NW
Washington DC 20554

MM Docket No. 87-286

Dear Sir:

Enclosed is the original and six copies of Part II of my comments in response to the Fifth Further Notice of Proposed Rule Making in the cited docket. The first part of these comments was filed on 19 June 1996. I have also sent copies directly to the Commissioners.

Very truly yours,

William F. Schreiber

No. of Copies rec'd 0+6
List ABCDE

Phone: 617-253-2579 FAX: 617-253-7302 E-mail: wfs@image.mit.edu

THIS IS THE ORIGINAL

THIS IS THE ORIGINAL

THIS IS THE ORIGINAL

THIS IS THE ORIGINAL

THIS IS THE ORIGINAL

RECEIVED
JUL 11 1996
FCC MAIL ROOM

Before the Federal Communications Commission
Washington DC 20554

In the Matter of
Advanced Television Systems
and Their Impact upon the
Existing Television Broadcast Service

MM Docket 87-268
Fifth Further Notice of Proposed Rule Making

Part II: What Kind of DTV Standard Do We Need?

Comments of

William F. Schreiber
Senior Lecturer, Professor of Electrical Engineering, Emeritus
Research Laboratory of Electronics
Massachusetts Institute of Technology

Submitted 10 July 1996

The opinions in these comments are those of the author only.
He has no financial dealings with any computer company.
Since his retirement in 1990, the author has had no role in
directing MIT's Advanced Television Research Program.

36-545 MIT, Cambridge, Mass. 02139 617 253 2579 wfs@image.mit.edu

Executive Summary

After a 9-year process, the Advisory Committee on Advanced Television Service (ACATS) has recommended a standard for digital terrestrial broadcasting (DTV), developed by the Grand Alliance (GA). Assuming that the Commission still believes that a transition to DTV is in the public interest, a standard should be issued that is likely to facilitate reaching specific objectives formulated in accordance with the Commission's statutory responsibilities. Such objectives might include the rapid penetration of the new system so that NTSC can be turned off after a reasonable transition period, the protection of the investment that must be made by industry and viewers alike, and the encouragement of nondisruptive improvement over time. Such improvement should include new services involving computers and communication networks such as the Internet.

The standard must be sufficiently detailed so that all stakeholders will have confidence that equipment built and products purchased in accordance with the standard in fact will work together initially and will continue to work as the system is upgraded. Digital systems like that of the GA are inherently much more complex than analog systems such as NTSC, so that it is to be expected that the standard must be more complex as well.

If it works as generally expected, the GA system is likely to deliver benefits to the public and to industry fully commensurate with the needed expenditures. However, it is not perfect. It includes no migration path for the anticipated improvements, there is no provision for receivers for less-demanding applications as cheap as now available for NTSC, it has less than the maximum possible spectrum efficiency, and it has limited interoperability and aspect ratio. Its most serious defect is that it permits the use of interlaced transmission formats. Since the use of interlace would confer no benefit of any kind, monetary or otherwise, on any domestic stakeholder, and since its use is likely to inhibit any nondisruptive improvement over time, it has no place in any new TV system. The DTV standard should therefore include only progressive transmission.

Certain changes might be made to the GA proposal that would make it more flexible, more attractive for non-TV imaging applications, and yet not significantly raise the cost of receivers or broadcasting equipment. These include allowing any frame rate (fps) that is a multiple of 12 fps, and any spatial resolution that is a multiple of 144x256. It is also likely that the standard can be simplified to some extent.

Although ACATS has done its job well, it is hampered by the fact that almost all the participants work for companies with a financial interest in the outcome. The public has been inadequately represented. To make the final desirable modifications of the proposed standard, the Commission is urged to appoint a small panel of independent experts, including FCC staff, with the ultimate decisions being made by the Commission itself.

1. Is a Standard Needed at this Time?

The Commission is at the end of a nine-year process that was intended to begin the transition from the NTSC system to an entirely new TV broadcasting regime that promises greater efficiency in the use of the spectrum, improved picture and sound quality, and, potentially, a host of new services. The cost of the switchover will be fully justified if the hoped-for results are actually achieved. However, this cost will be considerable to all the stakeholders, including manufacturers, broadcasters, program producers, and the viewing public, as well as to imaging industries other than those devoted exclusively to television. It therefore behooves all involved to think carefully about the probable effect of the pending decisions on the course of events.

The investments required to establish a new system are unlikely to be made unless everyone involved has confidence that the new system will be stable for a long-enough period of time. To ensure that stability, a standard must be set with sufficient detail so that the equipment initially installed by broadcasters and viewers alike will continue to operate successfully as the expected further development proceeds. Once enough receivers are in place to provide a viable market for commercially sponsored broadcasts, the need for a detailed standard will diminish. However, it is absolutely essential at the outset of service.

An example of what can happen when a new service is introduced without agreed-upon standards is satellite broadcasting in the US. This has gained viewers at an unexpectedly fast pace. About 3 million set-top boxes have been manufactured, and they are not interchangeable even for their intended purpose. In addition, none of them does either progressive scan or HDTV.

The signal produced by the GA system is not describable in a page or two as was possible with NTSC. Its structure is complex and there are many parameters that can be selected at will by the broadcaster. Barring the regulation (or even labelling) of receivers, the signal characteristics must be described in sufficient detail so that any receiver or set-top box that can deal properly with any one legal GA signal can deal with all legal GA signals.

A final consideration is that the current Inquiry has been under way since 1987, under the general oversight of the Commission and guided by a set of carefully prepared NPRMs. The work was done, for the most part, by employees of the TV industry. The process has occupied the attention of that industry during this period, and received a good deal of press coverage. Although the total amount of money so far spent is less than Detroit spends to put out a single new car, it is considerable in the minds of industry. After all this time and effort, a statement by the Commission that no new standard is needed would be greeted with dismay. It would make it very difficult to carry out a similar process in the future. In effect, the Commission has asked the industry to develop a new standard, and the industry has complied. The Commission should therefore issue a new standard. However, it would be an abdication of its duties, and especially of its duty to act in the public interest, to fail to scrutinize the proposed standard with great care to ensure that it will promote appropriate objectives in accord with the Commission's statutory responsibilities.

2. Objectives of a New TV Standard

There is no point in renovating the existing TV broadcasting system, which is both popular and profitable, without some specific goals in mind. The general goals are quite clear from the Commission's earlier decisions. We also want the new system to be so attractive to the stakeholders that they will go along with the eventual shut-down of NTSC. In this list, I have pointed out what seem to me to be some essential components of the overall goal.

2.1 Encourage the rapid penetration of the new system and discourage the proliferation of noncompliant receivers for cable, DBS, and niche markets. The political feasibility of the phaseout of NTSC, which is highly desirable just for the purpose of making more spectrum available for new mobile services, depends on the existence of a goodly population of DTV receivers as well as on the availability of inexpensive set-top converters. The standard should be shaped in such a way that receivers for any new TV service can be used for all new TV services.

2.2 Protect the investment of the public and the industry. The new receivers and broadcasting equipment should be usable for at least 15-20 years. Note that monochrome NTSC has been with us since 1941, and the color standard has remained essentially unchanged since 1953. The standard should, as far as possible, avoid the premature obsolescence of the initially installed receivers and equipment.

2.3 Enable nondisruptive improvement over time. It is agreed by most parties that the system should eventually provide progressively scanned imagery of at least 1000 lines vertical resolution. It is also widely expected that a host of new services and characteristics will become available, such as interactive programming, commercial activities, video on demand (VOD and NVOD), and Internet access. A migration path to such improvements, however, has not been specified. In order to ensure that the changes really are nondisruptive, the outlining of at least some possibilities would seem to be called for at the outset.

3. Is the GA Standard the Right Standard?

The criticisms made by the computer industry are not the only possible objections to the GA proposal. It is important to point these out, since some may be dealt with by small changes to the standard.

3.1 No migration path. Although lip service is given to moving to progressive scan eventually, and possibly to higher resolution, there is nothing in the proposed specification that is helpful in this regard. Some compatible improvements can be made at the encoder, such as better motion estimation. However, the implication in much of the talk about migration is that we shall somehow learn how to achieve a higher compression ratio. For that avenue, some requirement must be in the standard so that the early receivers can deal with the changed signal. On the other hand, if migration is to be carried out by sending a second signal for purposes of enhancement, the market will guarantee that existing receivers will continue to be served if there are enough of them in use before any enhancements are introduced. [1]

3.2. The use of interlace. There is universal agreement that progressive scan provides higher quality and easier interoperability, but some still advocate permitting interlace as an "interim" measure because of some supposed advantages, such as cost or resolution, or the lack of a good progressive-scan camera. In my previous submissions, I have disposed of all of these arguments, and I have also explained the danger that progressive scan will never be used if interlaced transmission is permitted. With the introduction of a high-quality high-sensitivity progressive HDTV camera by Polaroid, there is now no advantage of any kind to any domestic stakeholder, monetary or otherwise, in using anything other than progressive scan from the outset. [2]

One point that appears not to have been fully appreciated by some interlace advocates is that a progressive transmission with the same number of lines per frame as an interlaced transmission, and therefore with twice the analog bandwidth, can be transmitted at the same data rate as the interlaced transmission when MPEG-coded. In my Comments of 14 June to this NPRM, I included a number of papers from reputable laboratories (AT&T, NHK/RAI, and Thomson/UCL in France) attesting to this point. A more detailed paper on this subject is attached. This study, a part of the RACE/HAMLET project of the European Union, is by far the most detailed that has been done. It confirms the earlier finding that *progressive transmission, with its higher quality, complete freedom from interlace artifacts, and much better interoperability, does not require a higher coded data rate.* Given this fact, the proposal to use interlace is incomprehensible.

3.3 No provision for inexpensive receivers or set-top converters. Unfortunately, this is fundamental to the GA system, and cannot be fixed later. Every receiver, even the small one in the kitchen, must include a complete HDTV decoder in addition to all the components now required for NTSC. It will therefore be more expensive. No matter how long delayed the eventual abandonment of broadcast NTSC, there will still be a considerable number of NTSC receivers in use, especially since it seems unlikely that DTV receivers will ever be available at the same low prices as today's cheapest NTSC receivers.

NTSC receivers are also used for cable and DBS reception, as well as the playback of videotapes, of which an enormous number are in the hands of the public. The Commission has chosen not to regulate standards other than in terrestrial broadcasting, and therefore does not have the power to turn off NTSC in the alternative media. It is possible that engineering ingenuity and further development of integrated circuits can solve this problem. This issue is so important to the feasibility of the eventual termination of NTSC broadcasting that it might be appropriate for the Commission to obtain some independent professional opinions on its prospects before issuing new standards.

3.4 Less than maximum possible spectrum efficiency. While the GA system, if it operates as widely expected, will ultimately be able to provide about the current degree of choice of programs to viewers with half the spectrum allocation as NTSC, a cellular system, in which the coverage area of each station would be determined by the locations of a matrix of low-power transmitters, each emitting the same signal, would do substantially better. In such a system, each viewer could receive twenty different stations with an allocation of only twenty 6-MHz channels, as compared with today's 67 channels.

In order to operate in a cellular system, the channel coder and modulation method must provide a

very high degree of immunity to multipath distortion. This can be achieved with an OFDM system using frequency-domain equalization. In such a system, which can tolerate echoes as strong as the main signal, nondirectional antennas would suffice at almost all locations. [3] Unfortunately, the GA system, which uses a single-carrier system and time-domain equalization, would have to use very high-performance antennas at all locations in order to operate in the cellular mode. (In fact, even for the intended mode of operation, using a single high-power transmitter for each station, the GA system will require properly oriented antennas of rather good performance at almost all locations for satisfactory operation.)

3.5 Limited aspect ratio and interoperability. Computer interests have criticized both interlace and non-square pixels in the GA standard, while some Hollywood interests have called for the ability to operate in any aspect ratio, at least as high as 2:1. It is true that the best aspect ratio for a dramatic presentation is that in which it was first made. However, aspect ratios as high as 2.35:1 are in use, and it is simply impractical to make TV displays, particularly the hoped-for flat panels, of this shape. Letter-boxing, which is commonly used in Europe for showing wide-screen movies on TV, and is already used to some extent in the US, can readily be used with digital broadcasts. For example, the 2:1 shape desired by the film interests can be shown on a 16:9 display with a loss of only 1/9 of the screen height.

The desire for progressive scan and square pixels by computer interests is much easier to accommodate. In my opinion, these features are highly desirable in TV as well. As mentioned above and in my previous submissions, there is no valid reason whatsoever for using interlace in any new TV system. The same economies can be achieved with transmissions in progressive scan as in interlace, and with none of the bad effects. For example, an interlaced receiver, which would be of somewhat lower cost and of noticeably lower quality than a progressive receiver, can as well be used with progressive as with interlaced transmission. For the broadcaster, the cost of upconverting interlaced material to progressive scan will be negligible compared to the other costs of converting to DTV.

Square pixels are less important to the computer industry than progressive scan. Some TV people favor a 480x704 format for SDTV because it requires somewhat less compression. It would not be the end of the world to permit this format. However, the rather small negative impact of the exclusive use of progressive scan transmissions with square pixels on certain TV operations should be weighed against the benefits of much easier transcoding from one format to another. Considering all the transcoding that is used in TV today, with more to come as the new digital formats are added, interoperability is as important to TV as to computers. The weight of the argument is on the side of square pixels.

4. Feasible Variations of the GA Standard

Many of those who have been involved in the Inquiry, and particularly those companies that hope to profit from the speedy introduction of DTV, have stated that it is too late to make any changes at all to the proposed standard. In my view, this is an unsupportable position. The new standards are the first major change in TV since 1953, and have not yet been put into use. There is, as yet, no archival material in this format. If the job is done right, the new system should last many decades. Its longevity depends, to some degree, on its acceptability to all the stakeholders. When serious

objections are raised by significant industries, attention must therefore be given.

Having studied this matter at length, I believe that there are a number of modifications that could be made that would deal with most of the objections and that would not injure the interests of any domestic stakeholder. Furthermore, it would not take long to decide on just what modifications ought to be made, so that the standard-setting process need not be unduly delayed.

4.1 Elimination of the interlaced transmission formats. This is the easiest and most important change to make. Since no one would actually be injured by this change (except, perhaps, manufacturers of 1125/60 interlaced production equipment), and since many beneficial results would flow from this change in the proposed standard, there seems no reason not to take this action. [2]

An important related issue is the continued ability to use the first receivers as the system is upgraded over time. (Note that an NTSC receiver manufactured in 1955 can still be used today.) This issue, which has been on the Commission's list of desiderata since early in the Inquiry, did not receive enough attention in the deliberations. The primary change that is needed in the GA proposal to ensure this desirable outcome is the elimination of all interlaced transmission formats, as this makes transcoding from the transmission format to the display format quite simple. This action should be supplemented by a statement of principle by the Commission that evolutionary changes in the system shall be made in such a way as not to alter the transmitted signal so as to make it unreceivable on initial receivers. The statement could be made more precise by listing certain receiver parameters, such as amount of memory and processing speed. The cost and capabilities of an MPEG decoder depend to a great extent on these particulars. Therefore, one would need to limit the spatial and temporal resolution of the possible formats to avoid inflicting unnecessary costs on initial receivers. Alternatively, one could list the parameters of the transmitted signal that would imply corresponding receiver capabilities.

4.2 A range of frame rates. The GA standard calls for rates of 24, 30, and 60 frames/sec (fps). The cost of the receiver would not be significantly increased by permitting rates of any multiple of 12 fps up to some reasonable limit. Most receivers would have only one display rate, of course. This change would make it practical to use a display rate of 72 fps, which would help prevent flicker on wide, bright displays. The cheapest receivers would convert from the transmission rate to the display rate by dropping and repeating frames. High-end receivers would use motion-compensated interpolation, and might have both 60 and 72 fps display rates. [4]

Improvement in motion estimation and motion-compensated interpolation may permit the transmission frame rate to be reduced, enabling higher spatial resolution while maintaining good motion rendition. One candidate frame rate is 36 fps, which would make the 72 fps display very convenient, and might well eventually provide motion rendition as good as now seen in NTSC.

An alternate approach would be to require all transmissions to be at 24 fps, progressively scanned. This would make low-end receivers cheaper than with the GA scheme, but would require some additional expense for broadcasters, since existing NTSC programming would have to be transcoded. (Note that a substantial proportion of the "NTSC" archive has been converted from 24-fps film, and can be reconverted very easily.) In the case of video-originated NTSC material, motion rendition

would not be as good on the cheapest receivers, but could be made comparable, even with the knowledge we have now, in high-end receivers.

4.3 A range of spatial resolutions. The proposed GA standard calls for several different resolutions from 480 to 1080 lines/frame. It would not materially increase the cost of the cheapest receivers to permit many more possible resolutions. For example, if only the 16x9 transmission aspect ratio were to be used, then any resolution that is an integral multiple of 144x256 could be allowed, up to some reasonable limit. (The common factor of 16 is for MPEG convenience.) If the aspect ratio were permitted to vary, then resolutions of 16Nx16M, where N and M are integers, could be allowed, again up to a limit. Naturally, the signal must be labelled so that the receiver can determine which resolution is in use. Most receivers would have only one display resolution, and, of course, the more expensive the receiver, the higher this would be. [4] The ability to subsample/interpolate to the display resolution from the transmission resolution is already at least implied by the GA standard. In any event, with progressive scan, such interpolation would be of negligible cost.

4.4 Simplification of the standard. Some objections have been raised to the large amount of detail in the standard. It is certainly true that much more detail is present than in the case of NTSC, which can be described in a few pages. It should be borne in mind that the GA system is extremely complex. In terms of sheer computational power, a DTV receiver will be by far the most complex device in most homes. Ironically, much of the complexity of MPEG-2 is the result of herculean efforts on the part of the designers to obtain high compression with interlaced material.

It is inherent in all-digital transmission, in which the decoded signal is merely a string of ones and zeros, that the meaning of each bit of information must be accurately specified. The packet structure, while advantageous in many respects, calls for a lot of explanation. In addition, a great deal of flexibility is built into the source coder, and the receiver must be able to decipher the control information as well as to react properly to it.

There are two reasons for detailed standards. One is to ensure that all manufacturers initially produce compatible products so that DTV may make a rapid penetration of the market. The other is to ensure that improvements can be introduced without making the initial equipment obsolete. To do this, the format of the transmitted signal, together with its possible variations, must be described in sufficient detail to make the required receiver capabilities plain. Alternatively, the receiver capabilities can be delineated, and it can be left to manufacturers and broadcasters to produce signals that can be properly received. In the past, the Commission has greatly preferred to regulate transmission formats than receiver properties. However, it is well to keep in mind the great success of the All-Channel Receiver Law. At no significant cost to anyone, this law made the use of the UHF spectrum commercially feasible. It is possible that some combination of receiver regulation and signal specifications would prove optimum.

5. A Process for Modifying the Proposed Standard

It is evident from the foregoing that certain modifications to the proposed DTV standard would make it more attractive to a wider group of stakeholders and that such modification is feasible without injuring any domestic interest. A procedure is required for deciding on the appropriate changes so as

to meet the Commission's objectives and, at the same time, retaining the enthusiasm of broadcasters and equipment manufacturers.

The Advisory Committee, in most respects, has served the Inquiry well under the able leadership of Richard E. Wiley. Its main limitation is that it is composed, for the most part, of representatives of companies with a financial interest in the outcome. The public interest has not been sufficiently represented. Now that there is a fundamental disagreement between the TV industry and the computer industry over certain aspects of the standard, ACATS is no longer a suitable venue for making the final decisions. A small panel, appointed by the Commission, and composed exclusively of persons with no financial interest in the outcome, would seem appropriate. Business knowledge of both industries is required as well as technical expertise on the matters to be decided, not necessarily in the same individuals. Commission staff should be fully participating members of the panel, and decisions should be made by consensus. The terms of reference of the panel should be carefully formulated by the Commission, with the public interest at the top of the list.

6. Conclusions and Recommendations

After an Inquiry that has lasted much longer than originally anticipated, the Advisory Committee has recommended a standard for digital television broadcasting. The proposed system is not the best that might be formulated. In particular, it will not achieve the maximum possible spectrum efficiency, it will require directional antennas at most receiver locations, and it has limited aspect ratio and interoperability. However, if the system operates as expected and if the benefits that are predicted actually materialize, it will represent a substantial improvement over NTSC in many respects, particularly the amount of spectrum that must be allocated for a given degree of program choice for viewers. I therefore recommend that the following actions be taken:

1. Issue a standard for the format of the transmitted signal that is substantially in accordance with the proposal.

2. Eliminate from the standard all formats that do not use progressive-scan transmission and square pixels.

3. Permit any spatial resolution in multiples of 144x256 for the 16x9 aspect ratio and spatial dimensions of arbitrary multiples of 16 for other aspect ratios, both up to some reasonable limit.

4. Permit any frame rate in multiples of 12 frames/sec, up to some reasonable limit.

5. Direct that system improvements be compatible with receivers suitable for the initial system or else operate by transmitting an enhancement signal (or signals) in addition to the signal directed to the initial receivers.

6. Appoint a panel of experts having no financial interest in the DTV standards situation, and request recommendations and comments on the following points within a 6-month period:

   a. Possible simplification of the standard.
   b. A minimum-regulation strategy for both the transmitted signal and for receivers to ensure that initial equipment and receivers may continue to be used as the system is upgraded over time.
   c. The feasibility of Recommendations 3 and 4, and the technical details of Recommendation 5.

Notes

1. Enhancement signals might be sent in separate channels, but this is unlikely. A more realistic method is to superimpose a second data stream onto the main signal by splitting each point in the transmission constellation into a small group of four points, thus adding 2 bits per symbol. The extra information would appear as a small amount of random noise to the original receivers, but would be recoverable by high-end receivers that had a higher SNR because they used a superior antenna or were located closer to the transmitter. (In a VSB signal, each level would be split into a closely spaced pair of levels, adding one bit per sample.)

2. A detailed discussion of the pros and cons of interlace vs progressive scan can be found in my Informal Reply Comments to the Fourth NPRM, submitted 11 March 1996 and my earlier Comments to the Fifth NPRM, submitted 14 June 1996.

3. Such a system is described in the papers included in the Appendix of my Comments to the Fourth NPRM, submitted 9 August 1995.

4. Displays using cathode-ray tubes are subject to added cost if the horizontal sweep rate is adjustable. However, it is easy to change the number of samples/line, lines/frame and the number of frames/second, thus changing the spatial resolution. Flat panel displays must have a fixed spatial resolution, but the number of frames/sec can readily be changed, within limits.

Appendix

P. Guillotel and S. Pigeon, "Coding Efficiency Comparisons of both Interlaced and Progressive Scanning Formats." This study, a part of the RACE/HAMLET project of the European Union, is by far the most detailed study ever made comparing the coding efficiency of these two formats. I and P sources, I and P transmission, and I and P display were studied for a number of different subjects. Assessment was by SNR measurements, relative compressibility, and subjective quality. The general conclusion is that a progressive video signal with the same number of lines/frame as an analog signal, and with the same resolution, in spite of the fact that it has twice the analog bandwidth, can be MPEG-coded using the same data rate. It thereby confirms the conclusions in the earlier papers from Bell Laboratories and NHK/RAI that were included with my submission of 14 June 1996.

Project Number :	RACE 2110
Project Title :	HAMLET

CEC Deliverable N°. :	R2110/WP2/DS/R/012/b1
Contractual Date of Delivery to CEC :	Y2/M12
Actual Date of Delivery :	Y2/M12 (December 1995)
Title of Deliverable :	**Coding Efficiency Comparisons of both Interlaced and Progressive Scanning Formats**
Workpackage involved :	WP2
Nature :	Report
Authors :	P. Guillotel (Thomson Broadcast Systems) S. Pigeon (UCL)

Abstract. :

The purpose of this deliverable is to compare the coding efficiency of both interlaced and progressive scanning formats by means of MPEG-2 simulations. Statistical results about frequency homogeneity and motion compensation performances are first presented to evaluate how much a format can be compressed compared to the other one. A second part is dedicated to the coding efficiency itself by means of MPEG-2 encoding considering the same channel bit-rate whatever the transmission format.

Keywords

Scanning Formats, Progressive, Interlaced, Deinterlacing, Interlacing, Coding Efficiency, MPEG-2 Encoding, DCT Distribution, Motion Compensation, Bit-rate Control.

Summary

Interlaced versus progressive scanning is an important issue when dealing with digital television. Not only because the change from analog to digital communication may be seen as an opportunity to change formats, but also because of the well-known artifacts of interlaced scanning (interline twitter, line crawling, and field aliasing) compared to the natural way of representing two-dimensional images as the progressive format does. However, digital broadcasting has to face the problem of transmitting twice the bit-rate of the progressive format. It is the purpose of this deliverable to study this problem, and especially to check if the increased vertical and temporal correlations of the progressive pictures provide a significant improvement in the coding efficiency. In that case, progressive scanning may also be used as an intermediate transmission format to improve the coding performances of interlaced sequences.
Moreover, the main criteria in digital TV encoding is the picture quality, then, assuming that progressive display is more pleasant than interlaced, subjective picture assessment should be taken into account to evaluate how much a progressive picture can be compressed compared to an interlaced one. To answer all these questions is the objective of the scanning formats extension of the RACE R2110 HAMLET project (High Definition Advanced Multilevel Encoding Techniques). Within the framework of the workpackage 2, two digital transmission chains have been simulated, an interlaced one and a progressive one by means of MPEG-2 MP@ML encoding, optimized for each format. Progressive and interlaced source sequences are used, deinterlaced and interlaced when necessary, to perform successful comparisons between both format in each possible configuration.

Contents

Summary 2

1 - Introduction 4

2 - Progressive Encoder 5

3 - Interlacing and Deinterlacing 6

4 - Statistical Properties of Both Interlaced and Progressive Formats 7

4.1 MPEG-2 Encoder Parameters 7

4.2 Source Sequences 8

4.3 DCT Coefficients Distribution 9

4.4 Motion Compensation Performances 12

4.5 Coding Gain 14

4.5.1 Interlaced versus Progressive *14*

4.5.2 Influence of the Increased Vertical Resolution *15*

5 - Coding Efficiency of Both Interlaced and Progressive Formats 18

5.1 MPEG-2 Encoding Parameters 20

5.2 Coding Efficiency Comparisons 20

5.2.1 PSNR Measures *20*

5.2.2 Subjective tests *27*

5.2.3 Bit-Rate Control Parameters *29*

5.3 Influence of the Bit-Rate 34

5.4 Influence of the Picture Complexity 34

5.5 Influence of the Deinterlacing 35

6 - Conclusion 36

References 37

Coding Efficiency of Interlaced and Progressive Scanning Formats

1 - Introduction

It is not the purpose of this paper to cover the historical considerations that led to the adoption of an interlaced scanning television, or to discuss the well-known interlaced artifacts described in [1], but rather to investigate new techniques to improve the efficiency of MPEG-2 based-encoders, and especially, within the HAMLET/WP2 project, to study the influence of the scanning mode on the coding efficiency and the possible use of the progressive as an intermediate coding format [2]. Three deliverables are planned to sum up all the studies performed on that subject, the first one is a list of theoretical and technical considerations on the advantages and drawbacks of interlaced and progressive scanning formats [1], the second specifies a generic format converter [3], and the last is the present one.
This paper originates from the following comments : progressive format seems much more attractive than interlace for signal operations, progressive format has the advantage of compatibility with computer graphics, multimedia applications and film production, and finally progressive scanning allows a better display. Unfortunately, progressive requires twice the number of pels of interlaced signals, and is more hardware consuming. Despite these drawbacks progressive is very attractive, and various results are already available [4,5,6,7], but they differ on their conclusions. Thus, further investigation is required, which is the object of this deliverable.

Both progressive and interlaced source pictures have been used for the simulations, because of the influence of the original format. Progressive to interlace and interlace to progressive conversions insure the compatibility between each format. Simulation results should allow the following questions to be answered : which kind of scanning format must be used for input images ? What is the most adapted display mode ? And finally what is the best transmission format ?

After a presentation of the progressive MPEG-2 encoders used for testing, and especially the simplifications done, a chapter is dedicated to the statistical properties of both formats in order to evaluate how much they can be compressed regarding their spatial homogeneity and motion interpolation capabilities (thanks to their temporal correlation). This chapter is followed by the main part, which contains the encoding simulation results together with subjective picture evaluation carried on by expert viewers.

Based on these results different conclusions are drawn to show that progressive display improves the overall picture quality, without loss of coding performances compared to the existing interlaced format. And finally, it leads to the statement that progressive transmission may be also an intermediate step towards progressive broadcasting.

2 - Progressive Encoder

A progressive MPEG-2 software encoder has some parameters to be set to be MPEG-2 compliant. Some are specific to the progressive format and optimized for it [8]. The MP@ML profile is used, even if 625/50/1 can not be transmitted using it. This is firstly because the objective of this deliverable is to compare both formats with the same picture size, and secondly because a new level might be further included in the MPEG-2 final standard specification to comply with progressive scanning.
Within HAMLET/WP2 two software encoders have been developed and will be called Thomson scheme and UCL scheme [5] in the following (the default scheme is the Thomson's one). They differ mainly by the motion estimator and bit-rate control.

The parameters specific to the progressive format are the following :

- *progressive_frame* set to 1, coded video contains only progressive frame pictures. In this case, restrictions apply as follow : *picture_structure* = "frame" and *frame_pred_frame_dct* = 1, allowing only frame DCT and frame predictions to be used ;

- *frame_pred_frame_dct* set to 1, affects the syntax of the bitstream. For each macroblock, this flag suppresses useless flags like *frame_motion_type* (2 bits) and *dct_type* (1 bit) from the bitstream;

Accordingly, progressive coding allows to reduce the side-information by 3 bits/macroblock. It also lowers the number of vector to be transmitted since no field motion vectors exist in such case.

Besides the MPEG-2 syntax, the new sampling grid structure of the progressive format allows some possible simplifications :

- The motion estimation is based on a pyramidal structure which leads to a very simplified and efficient data processing. Only 1 vector has to be computed (frame vector) when 5 are required for the interlaced format (4 field vectors and one frame vector) for only one temporal direction. Furthermore it leads to a simplified mode decision process.

- Chrominance filters for up and down sampling are less complex than in the interlaced case, for instance the following implementation has been simulated :

  ❒ Down sampling with one 5-tap FIR filter instead of two 7-tap FIR filters (one for each field);

  ❒ Up sampling with two 2-tap filters (one for each line parity) instead of two 2-tap and two 3-tap filters (one for each field and each line parity);

Other MPEG-2 parameters are the VLC intra tables *(intra_vlc_format = 1)*, the non-intra quantization matrix (flat), the macroblock mode selection, the thresholding of the DCT coefficients, the quantizer type (*q_scale_type* = 0), the zig-zag matrix (*alternate_scan*= 0). The values are the same for both formats for sake of simplification and because no significant difference has been observed with other choices. All these points are not in the scope of this paper and will not be further discussed.

3 - Interlacing and Deinterlacing

For the simulations, four different processes are used for interlaced to progressive and progressive to interlaced conversions. Whatever the format, these converters are key points in a broadcasting chain [9], therefore a special attention has been paid to this problem, and this is the subject of this chapter.

First of all, a distinction has to be done between process at the encoder side and at the decoder side. Whereas the first one requires a high picture quality, the second one can not use expensive tools. Having in mind these considerations, the encoder interlaced to progressive conversion is a high quality motion compensated deinterlacer from UCL [3,10], whereas the decoder one is a low cost macroblock-based motion compensated deinterlacer making use of the MPEG2 transmitted motion vectors.
The first application requires a finely tuned calculation of the motion vectors in order to recover, in the reconstructed progressive sequence, the initial temporal and vertical correlation existing inside the original "analog" signal. Once these vectors have been found, a deinterlacing method that handles field aliasing properly has to be used. This is achieved by means of the general sampling theory which was proposed recently to handle interlaced images and proved to be successful. Linear spatio-temporal interpolation has also been studied in [11], but complementary results are required before concluding.
The second deinterlacer is a low cost solution thanks to the absence of motion estimation. The different macroblocks make use of the MPEG-2 transmitted vectors when available, otherwise vertical interpolation is performed (for instance for intra macroblocks, intra frames). Motion compensated interpolation is fulfilled only for the luminance component, whereas the chrominance is vertically interpolated with a linear filter. Moreover, all these processes are macroblock based, which means that they do not require information from the surrounding macroblocks. It will be shown later on that this last requirement is too hard since a line structure appears in the borders of the macroblocks, thus better results can be expected with the neighboring blocks.

Concerning the progressive to interlaced conversion, it is performed by means of a simple 11-tap filter known as "HHI filter" (figure 1) followed by a vertical subsampling by a factor of two. This filter introduces a vertical definition loss according to the Kell factor. This factor is meant to reduce the line flicker that appears on bright and sharp horizontal edges when displaying in an interlaced format. The moderate complexity and sufficient efficiency of this filter make it suitable for both encoder and decoder applications.

-4	8	25	-123	230	728	230	-123	25	8	-4

Fig. 1 - "HHI filter" coefficients

4 - Statistical Properties of Both Interlaced and Progressive Formats

Three different statistical measurements have been performed to compare the potential ability of each format to be compressed. The first one is based on the frequency homogeneity (with the DCT coefficients distribution), the second one on the motion estimation behavior provided by the motion compensated DFD (Displaced Frame Difference) between the reference and interpolated frames, and the last one on the coding gain of progressive over interlaced scanning (without bit-rate control, i.e. with the same quantizer step size)

4.1 MPEG-2 Encoder Parameters

The aim of that section is to give (or anyway try to) a ratio value between bit-rates needed for a progressive transmission versus an interlaced transmission considering the same picture quality. Moreover, it will be interesting to have these values for each picture type (I, B or P).
To do so is quite difficult between different scanning formats, how can it be guaranteed to have the same picture quality ? How can it be guaranteed to compare exactly the same information ? However, the following parameters have been selected, even if other solutions might be used, having in mind that the progressive format performs probably worse than interlaced (in terms of compression rate when the quality is fixed), and that these values have preferably to be done is the worst conditions to obtain a low anchor :

-The encoder complies with the MP@ML profile of the MPEG-2 standard except for its use of the progressive 625/50/1 format (not currently supported by it);
-The picture quality can be considered at least as good when the same quantization step is used for each pictures and each format (worst case for progressive). In this section quality is related to the quantization step (the lower is the quantization step, the higher is the quality). It makes the comparisons between both formats more convenient and allows to work at "fixed quality". However, this definition does not take into account the properties of the human vision. The subjective picture quality of both formats will be discussed later on;
-The GOP structure is selected starting from the following point : assuming the classical GOP structure for interlaced I, B or P frame pictures (figure 2-a).

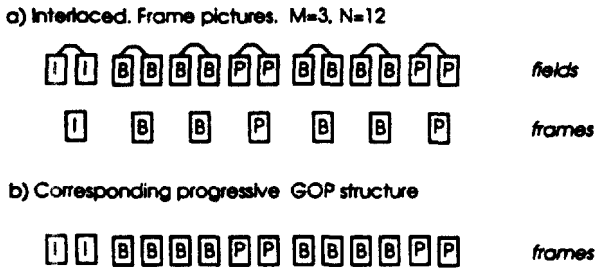

Fig. 2 - GOP structures used

After deinterlacing the same frames lead to two progressive pictures, corresponding to figure 2-b, the picture content between the interlaced and deinterlaced frames is thus exactly the same as well as the picture type.

Extrapolating this structure to the progressive source sequences leads to the non MPEG-2 GOP structure of figure 2-b. Of course, this structure is non-optimal for progressive scanning (again worst case), but it makes the comparisons between interlace and progressive easier in this first study. Another structure will be considered later on this deliverable.

With that trick it will be possible to compute directly for each picture the ratio between the bit-rate required for progressive coding versus interlaced coding when interlaced and progressive pictures are transmitted with the same picture quality.

It should also be pointed out that the reference I or P pictures for motion estimation and compensation are always the nearest ones as explained in Figure 3.

Finally the motion vector range is taken equal to [-63, +64] in the vertical direction, whereas [-127, +128] is selected for the horizontal one, whatever the sequence and the temporal distance between the reference and decoded pictures. These vectors are computed with a 5 levels hierarchical block-matching algorithm, and a search window set to (4×2) pels at each level.

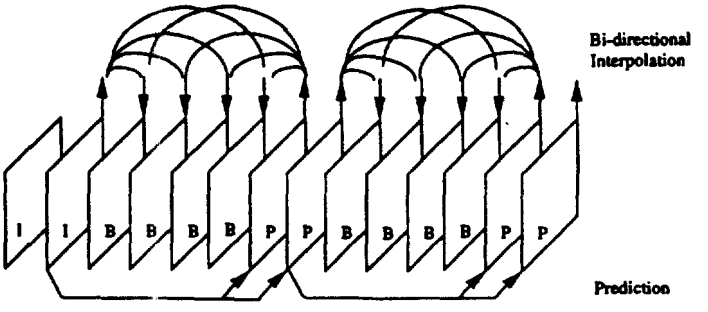

Fig. 3 - Reference pictures for progressive coding

4.2 Source Sequences

Four 720x576 50Hz sequences have been selected for the simulations. Two are interlaced and have been deinterlaced with a high quality motion compensated deinterlacer [3], and two are progressive sequences. The following gives their characteristics :

- **Interlaced :**

  # *Mobile and Calendar* : 50 interlaced pictures originated from a tube camera, its progressive version is obtained by motion compensated deinterlacing [3];
  # *Flower and Garden :* 50 interlaced pictures originated from a tube camera, its progressive version is obtained by motion compensated deinterlacing [3];

- **Progressive :**

  # *Renata RAI* : 100 progressive pictures originated from an HDTV tube camera, its interlaced version is obtained through vertical filtering (with the HHI 11-tap filter, taking into account the Kell Factor) and subsampling. In the following it will be called *Renata* for the progressive version and *Renata F* for the interlaced version (because of the Kell filter). Once it will be called *Renata F/F*, it means that both the interlaced output and the progressive input sequences are filtered.

Kiel Harbour : 100 progressive synthetic pictures, obtained by digitizing and processing a photo with synthetic motion, its interlaced version is obtained through vertical filtering (as for *Renata*) and subsampling. In the following it will be called *Kiel;* and *Kiel F* for the interlaced version.
Pendel : 50 progressive pictures, originated from a progressive tube camera;
Pops : 60 progressive pictures, originated from a progressive CCD camera;
Foot : 50 progressive pictures, originated from a progressive tube camera;
Kiel 2 : 40 progressive pictures, from the same sequence as previous Kiel but not at the same moment;

The first four sequences have been processed with the Thomson scheme whereas the last four are processed with the UCL scheme.

4.3 DCT Coefficients Distribution

An MPEG-2 encoder can process interlaced pictures in different ways. For instance a macroblock can be frame DCT coded or field DCT coded (the mode selection differs from the coding schemes), whereas only frame DCT is allowed for progressive pictures. In figure 4 and 5 the DCT distributions for these different modes are plotted, they represent the distribution of the squared DCT coefficients for the luminance component, averaged over all the DCT blocks (8x8) in each mode, for the whole sequence. In addition to the DCT mode, the distinction is done between intra (Figure 4) and inter (Figure 5) macroblocks.

One of the assessments at the beginning of this project was that the double bit-rate of the progressive scanning could be compensated by a better spatial and temporal correlation. This chapter tries to give some insight for this, and the conclusion depends on the scanning format and the amount of motion within the sequences :

1)- For sequences with motion (*Flower* and *Renata*), field DCT is known to perform better because of the absence of vertical frequencies (visible in the frame DCT mode). In that case progressive is expected to be better than interlaced when both formats have the same vertical resolution, and not worse with a double resolution since progressive is supposed to be more homogeneous. This is clearly what figure 4 shows;
2)- For sequences without motion (*Mobile*) frame DCT takes more advantage of the spatial homogeneity, and therefore is similar to progressive encoding. Thus no differences between both formats are expected and visible in figure 4.

The general behavior for both formats seems to be the following one. When the source sequences are interlaced, fixed pictures are similar for both scanning modes, and moving ones are better when deinterlaced. Extrapolating these results to the progressive case (twice the vertical resolution) leads to similar performances for progressive and interlaced coding for moving images, and better interlaced coding for non moving ones (to be confirmed).
Concerning the temporal correlation, it will be studied thoroughly in the next paragraph, because it is mainly linked to the motion estimator performances, to the deinterlacing efficiency and to the picture quality (noise).

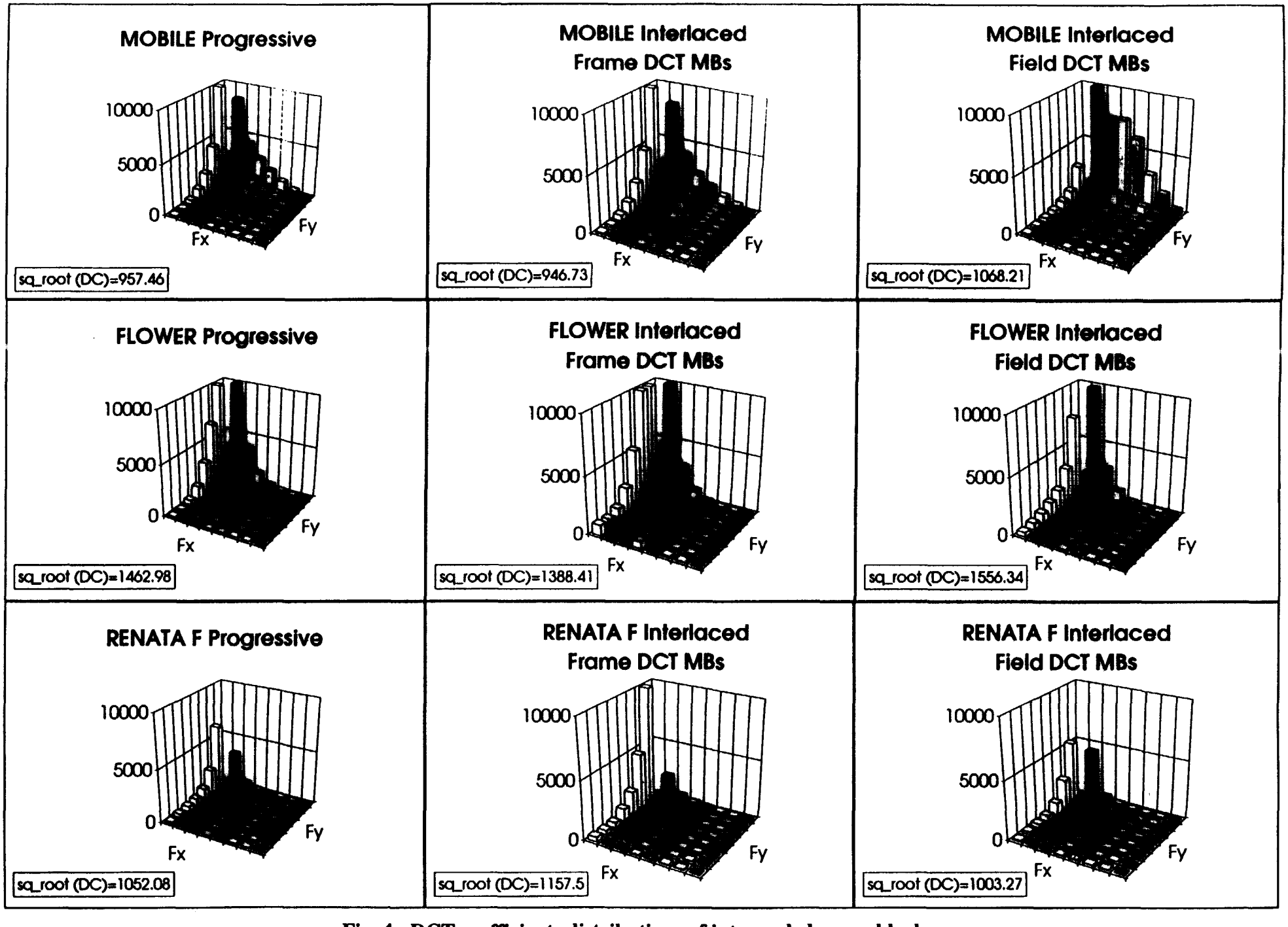

Fig. 4 - DCT coefficients distributions of intra coded macroblocks

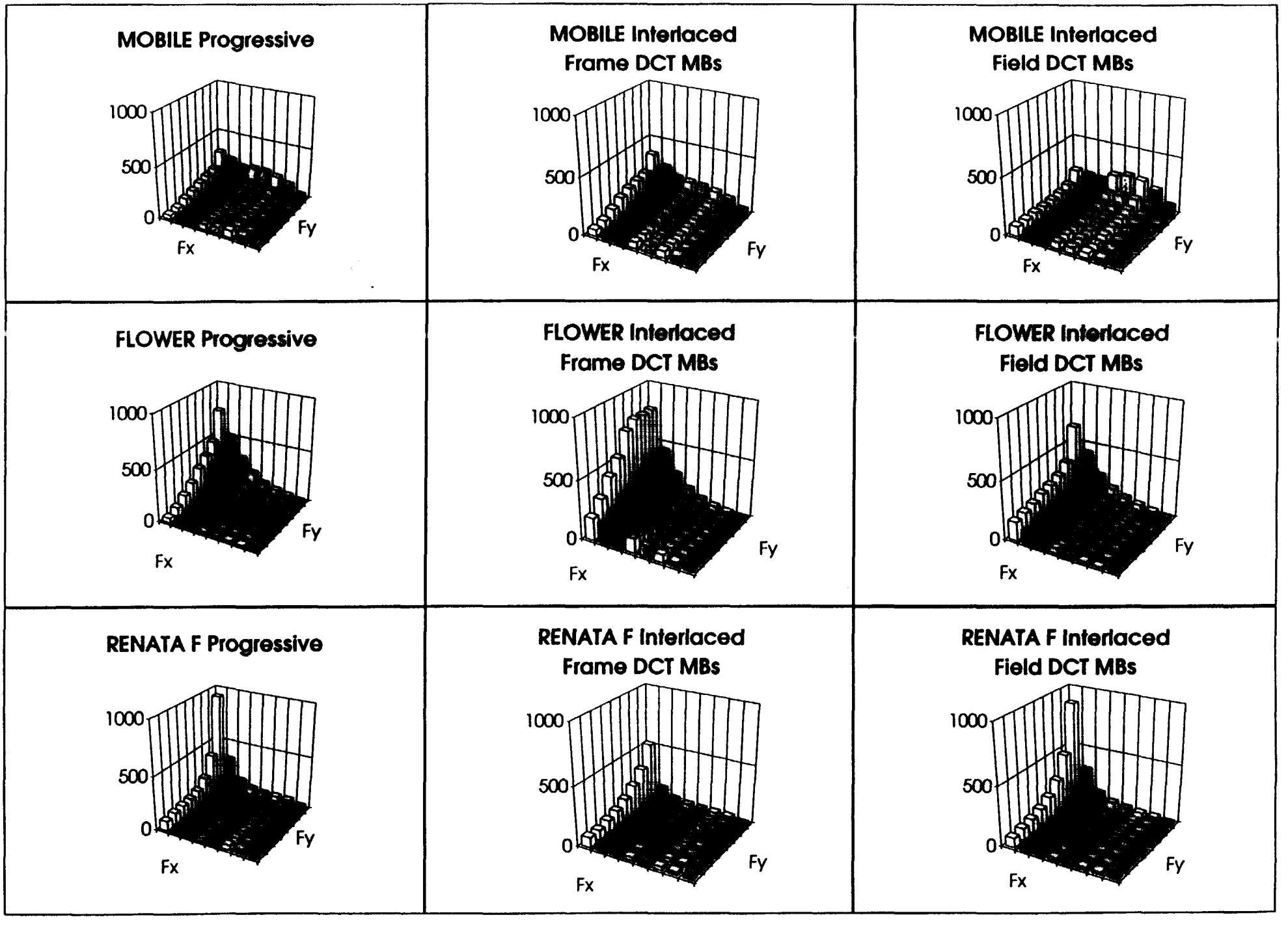

Fig. 5 - DCT coefficients distributions of inter coded macroblocks

4.4 Motion Compensation Performances

Another assessment concerning progressive scanning is that it improves the motion estimation performances. The MPEG-2 standard allows different kinds of motion prediction, either field or frame prediction, and forward (P frames) or bidirectional (B frames) in a macroblock basis. To measure the performances of a specific motion estimator, the prediction error can be used, it is also called Displaced Frame Difference (DFD). In Figure 6 the mean DFD values for the four first sequences are plotted for both interlaced and progressive processing, and for the P and B frames. Only the best prediction mode selected is used.
The unavoidable drawback of this trial is that the prediction error is computed between the current coded picture and the previous decoded one. The DFD depends therefore on the decoded pictures quality, and thus on the bit-rate. By using the same quantization step, this drawback is removed.

The conclusions from these figures are picture dependent as well as motion dependent as for the previous DCT distributions :

- ***Interlaced original pictures*** :

  - 1)- Without motion (*Mobile*) the picture is mainly frame coded, progressive and interlaced formats are thus similar. No differences are therefore expected except from the deinterlacing artifacts. The B frames are slightly better predicted in the interlaced case, and the P frames slightly better in the progressive one. It is probably due to the prediction structure (figure 3) which leads to a lower temporal distance for one of the two P frame;
  - 2)- With motion (*Flower*) the interlaced picture is mainly field coded, motion estimation is therefore difficult due to the high vertical frequencies. It leads to a more efficient progressive estimation helped by a deinterlacing without artifacts;

- ***Progressive original pictures*** : Interlacing performs better mainly because of the vertical resolution (Cf. the *progressive F* curve in the *Renata F* graph when the progressive sources are low-pass filtered). The greater resolution of the progressive pictures enhances the DFD measures mainly at the borders of moving objects. Another reason is the camera noise in the original progressive sequence (mainly in *Renata*) which explains that *Kiel* performs better than *Renata*.

The general behavior for both formats seems to be the following one, and confirms previous results from section 4.3 :
When the source pictures are interlaced, fixed pictures are similar for both scanning modes, and moving ones are better when deinterlaced. Extrapolating these results to the progressive case (twice the vertical resolution) and the help of the *Renata* sequence leads to similar or worse progressive coding for moving images (first half pictures of *Renata)*, and better interlaced coding for non moving ones (last half pictures of *Renata*).