INTERNATIONAL ORGANISATION FOR STANDARDISATION

ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 M9478
Pattaya, March 2003

Title: **Report on SynthesizedTexture CE**

Source: **Shlomo Birman (Vimatix), Ehud Spiegel (Vimatix) Cyril Concolato (ENST), Alexandre Cotarmanac'h (FT)**

Status **Core Experiment input**

Summary

This document presents the results of the SynthesizedTexture Core Experiment that was defined in N5286 and started after Shanghai meeting. After citing the core experiment purpose and workplan, we present the results and our conclusion.

Preliminary results presented at the MPEG#63 Awaji meeting showed that on average, VIM compression is better than BIFS in an average of only about 3 times. This called for improving the result.

This document present the proposed way of coding Synthesized Texture, which would achieve similar compression to VIM.

Background: SynthesizedTexture CE

The coding of SynthesizedTexture nodes seems to be under-optimized in BIFS compared to the proposed VIM stream (M8902). The purpose of this Core Experiment is to:

1) Confirm the compression results as presented by Vimatix;
2) Evaluate if BIFS results for SynthesizedTexture coding can be improved.

The steps in the CE are:

1) Confirm the compression results as presented by Vimatix against existing BIFS compression tool (predictive MF field). Vimatix will perform the tests with the support of Systems experts on how to use existing tools – 2 Week.
2) Recommend restructuring changes in ST nodes to better fit existing compression tools in BIFS – 1 week. ENST and France Télécom will propose restructuring changes.
3) Modify the ST nodes according to recommendations and provide new bitstream– 1 week. Vimatix.

4) Evaluate the modified ST nodes compression with BIFS compression tools (including predictive MF field) – 1 week. Vimatix.
5) Come up with final list of recommended changes that will result in the maximum gain with existing BIFS compression tools. Scene representation AHG.

Participants:

- Shlomo Birman (Vimatix)
- Cyril Concolato (ENST)
- Alexandre Cotarmanac'h (France Telecom)

Following the CE:

1) If the CE results a significant enhancement of the compression within existing BIFS compression tools, node restructuring will be formulated within an NB comment to the FDIS of amendment 4 for the next MPEG meeting. – Within 3 weeks.
2) If there's still inefficiency within BIFS, additional compression schemes within BIFS will be welcome.

Results presented at the Awaji MPEG#63:

The results of the CE have been presented in the Awaji MPEG#63 in m9237.

The clean-up of the file was done as follows:

- Fields with default values are not written.
- Floating-point fields with definitive limited set of values are replaced by integer fields.
- Fields having type-support for alternative options, but with no such alternative option implementation, are removed.
- SynthesizedTexture elements that are always at the same layer got a unified Z value. This means that these elements can be 2D instead of 3D and a single-value field is needed for the z order.
- Were applicable, elements of the same type are grouped together.
- The representation of patch elements previously used a mixture of different nodes, has been replaced with a much more efficient one using single new node only.
- The new Predictive MF field coding has been applied.

The results showed that on average, existing MPEG tools compresse now 4 times better than they did with the original nodes structure. At the moment, VIM compression is better than BIFS in an average of only about 3 times. Vimatix argued that these textures now carry reasonable file sizes with standard BIFS tools.

However, it has been decided to continue the CE in order to try and improve the results or suggest alternative ways. From w5316:

" The Systems subgroup acknowledges the results of the Core Experiment on Synthesized Texture. Between the Shanghai meeting and the Awaji meeting, redesign of that technology and better usage of the BIFS coding tools have resulted in 400%

compression gain improvement. There is evidence that an additional 300% gain is achievable. However, there is no consensus whether this additional gain (1) is achievable using the current BIFS framework, (2) is specific to a particular type of content, (3) resides in completely new technology that may lead to the creation of a new coding standard. The Systems subgroup therefore recommends delaying the promotion of the Synthesized Texture technology and investigating in the AHG on Scene Description which of the options listed above or combination of those is correct. In the meantime, the Systems subgroup recommends developing the Synthesized Texture technology within the on-going Advanced Text and Graphics amendment."

Work done following the Awaji MPEG#63 resolution:

Few suggestions for improving the compression under BIFS have been examined:

1. Restricting nodes data-type, since they will probably would not be used outside the scope of Synthesized Texture. This saves relative small amounts of data.

2. Since the high level of the structure in BIFS needs extra bits, reducing the amount of nodes from the same reason that they will probably would not be used outside the scope of Synthesized Texture. This can save, but by doing so we actually going towards the 'native' coding of a texture as a whole.

3. Removing un-used fields. This saves relative small amounts of data, and typically do not change the comparison results since typically they exist and 'waste' space in the native VIM coding as well.

4. Use indexes for the Width and the Brightness fields of the ColorProfiles. Using indexes for Brightness would not reduce almost the size, since on a typical (photo-realistic, not cartooned) image there are 5-20 thousands of ColorProfiles, altogether having 2-10 thousands of Brightness values, therefore the indexing process would not significantly reduce the amount of bits. Regarding Width, indeed in a small image there are relatively few values and a gain of few percents in file size can be achieved (not more than that since we already used a limited set of integers and the Predictive MS Field handled it very efficiently) ; however on the other hand it would significantly slow down the rendering and reduce the performance on mobile devices which are the target clients when referring to small image size.

5. Same argument as in 4 applies to the Patch node. Note that although in many cases the patch is similar to the background (only in color!), it needs full range since it can also carry significant 'noise'. We estimate that the Predictive MS field did well on patches were applicable, so the redundancy has already been compensated.

6. Using YcbCr instead of RGB and optimizing on the specific components. The problem is that his is mixing compression with rendering. The YcbCr in 'native' VIM coding is used to improve the compression, but for rendering we are going back to RGB. Yes, the

BIFS method could reuse the color space but in the expense of playing speed. Further more, as it is a quantization mechanism, it could be applied to all colors in BIFS, and it would be not such small issue to define that especially to ensure backward compatibility of streams.

7. Use of localization and occupancy grid as in 'native' VIM coding. The problem is that that this technique is global and considering simultaneously different types of image elements, in our case different nodes. Therefore, you could say that using this technique would be like taking a snapshot of a 2D BIFS scene and compress it. Doing so is out of the scope of this CE since it's a very complex technique, which should be developed to all BIFS image elements (and be matched to various scene types – for example, when not having many image elements, such as in simple Cartoons, it may be redundant or even adding bits)!

Conclusions and Next Steps:

The CE has shown that there is not much more to gain in file size by rearrangement and modification of the Synthesized Texture Nodes as it is defined now. However, there are two optional ways to achieve similar results as 'native' VIM coding:

1. Use a structure and coding similar to the 'native' VIM, but make sure to integrate it in the right way into the MPEG environment. This move involves also unifying the ST nodes into am efficient structure, which is further justifies since they will probably would not be used outside the scope of Synthesized Texture.

2. Trying to port some of the major efficiency contributors, such as items 6-7 above to the BIFS environment. This involves complex analysis to ensure the relative contributions, compensating for the use on only part of them, and implementing new BIFS techniques that should be also developed to match other uses, and tested and verified over many situations and Nodes combinations.

Since option 1 is proved to be working (see the 'native' VIM performance), and option 2 is out of the scope of this CE and would require a significant amount of work before even being able to define if it is applicable at all, and particularly to BIFS in general (if its only benefiting ST than it would be waste of time to implement while we have option 1 practically ready), our recommendations are as follows:

1. Start with option 1 above. ST will have a single Node/Point-of-reference, but with an optional switch to use alternate (BIFS-based) technique for coding. This would enter the CD in the Pattaya MPEG#64 meeting.

2. Carry a short CE until May, which will try to implement option 2 above. In case of a clear and fully implemented success, the switch can be used to direct to this alternative.

Vimatix will disclose its 'native' coding mechanisms enabling interested participants to learn it and suggest both improvements and ways of porting into BIFS-based structure. The coding mechanism is uploaded as contribution m9479. The new CE would be defined in the Patayya MPEG#64 meeting.

3. A bit outside of the scope of the CE: Upon carrying a long-range MPEG project which aims to improve BIFS coding techniques, Vimatix would contribute from its experience for the benefit of other MPEG tools.