Abstract – Image compression based on large blocks and multi-scale DCT transform is described. In this paper DCT-based schema is propose which provides quality comparable to or better than JPEG2000 at the same compression ratio. Support for efficient access to large images and region of interest coding are considered as native part of schema.

1. INTRODUCTION
Discrete Cosine Transform (DCT) is widely in use for image and video compression because of its near to the optimal ability to decorrelate signals. Research communities interest was later moved to wavelets because they promised better quality when compared to DCT. It appears to be true comparing JPEG [1] and JPEG2000 [1] standards, but such comparisons ignore other differences between standards, mainly entropy coding. It is well-known that block size in JPEG is not optimal, but no coding standard improved it.

Work on this compression schema began as part of digital microscopy project [2]. Considering size of even average microscope image it was realized that conventional formats offer direct access to part of image or compression, but not both, JPEG2000 is exception. It is unclear are available libraries written with support for fast access to large images. After a few initial tries with microscope images it was clear that this technique can provide high quality, good compression ratio (CR) and fast access to arbitrary part of image. Use of DCT is advantage considering post processing options like deblocking and deranging filters which are well documented. There is large number of very fast algorithms and circuits for calculating DCT, especially for small block which is another advantage.

It is unusual that there is not very much research literature about using DCT with larger blocks; exceptions are [3] and [4]. Some problems encountered with smaller blocks are absent and number of new problem arises. It is worth noting that comparing PSNR doesn’t provide even close correlation between quality because of completely different set of artefacts introduced by this technique and JPEG2000.

2. BASIC CODING
Basic coding is similar to classical JPEG. An image is divided into the blocks, each 64x64 pixels wide; each block is processed independently. This is somewhat different from results in [4], probably because used images are from microscope. Colour images are transformed from RGB to YCbCr colour space, reflecting fact that human visual system (HVS) is more sensitive to changes of luminance than of colour. DCT is applied to every block in all channels and quantisation is applied. To further exploit properties of HVS, channels are not quantised equally. Exact ratio of quantisation is open question, but for low to medium CRs luminance: chrominance ratio 1:2 seams to provide near optimal quality; for high and very high CRs situation is somewhat inverse, changing as CR increases to ratio 2:1.

Next step is to order block in zigzag sequence, perform run-length coding and some entropy coding. Choice of optimal entropy coding needs more research, especially considering additional levels of DCT. Using the technique mentioned above, it was possible to achieve compression levels and subjective quality near to JPEG2000 in the wide range of CR values, from 1:24 to 1:256. At high CR values in some aspects, quality was superior even with noticeable edges between blocks. Situations where those are true include texture-rich objects and small details; JPEG2000 blurs image resulting with loss of fine details on textures and even with loss of small shapes. In medical application it can be serious drawback and this is cited as serious problem with JPEG in [5]. For the highest quality compression threshold can be applied to coefficients.

3. MULTISCALE CODING
Obvious drawback of using large blocks and aggressive quantisation is removing of high frequencies from block’s spectra. That leads to blurred edges and ringing artefacts as it is inherent property of DCT and other Fourier-based representations. Similar to wavelet decomposition, DCT can be applied at different block size. In principle, idea is simple: take residual between original and compressed image, transform residual with DCT with smaller block and quantise it. Algorithm can be repeated iteratively, each time lowering block size. At first, it seems that nothing is gained in term of compression/quality ratio because for better quality additional bits must be allocated. Further analysis shows that smaller blocks have zero DCT coefficients everywhere where image is good represented with low frequency components and non-zero where distinct features like edges exist. Result is localisation of high-frequency features. In this research five block sizes are used: 64, 32, 16, 8 and 4 pixels.

4. IMPLEMENTATION OF MULTISCALE CODING
For maximal usage of spatial and frequency coherence luminance coefficients from smaller blocks are turned to zigzag sequence and grouped as in Table 1. Chrominance coefficients are interleaved and then grouped in same way. Resulting sequence maximise number of consecutive zeros, making run-length compression very efficient. To further increase compression, special symbol is used to mark zeros until end of appropriate part of bit stream. To simplify encoding and decoding and to improve compression, all symbols are 1 byte long; larger values are stored in 2 bytes (like DC value of block) but without special treatment and unused bits are sign extended. This ceased need for special symbols like bit length thus making bit stream more predictable and more entropy coder friendly.
5. ADVANCED FEATURES

There are few features which make this schema different from other, simpler approaches. Their usefulness depend primarily on area of use.

A. Access to arbitrary part of image

Considering size of single block before compression and after wide range of CRs it is clear that it can be completely coded alone without compromising CR. Further, its position in bit stream can be encoded resulting in small overhead. This provides fast access to arbitrary part of image. For faster reading decoder can convert offsets from relative to absolute which requires small amount of extra memory (for image 2048x2048, 24bpp it is only 4KB). For writing purposes, encoder needs only to encode changed part of image and to adjust offsets, without any need to perform encoding and decoding. Using above, even very large image can be encoded needing small amount of memory if they are provided in blocks of at least 64x64.

B. Region Of interest (ROI) aware compression

Block based schema naturally provides support for ROI compression. All blocks which are covered by ROI are less compressed. This rule can be applied to smaller blocks at all levels of compression. At the end of compression, ROI will be approximated with the smallest blocks. It is possible to include not just one ROI, but lot of them, each having its own quality factor. For decoder, this is almost completely transparent.

C. Decoding at smaller size

In some cases, user may desire to see whole image on screen, no matter how large is it. Brute force approach is to decompress it and then down sample to desired size. It is appropriate in cases when desired image size is a bit smaller then original. When there is considerable difference in image size, it is more economical to decompress image to nearest size available directly from compressed form and then to resample to desired size.

6. RESULTS

Prototype is implemented in MATLAB 7 environment with some functions from Image Processing Toolbox. Due space constraints, only the most interesting results are presented. Four pictures are used: Lena, Mandrill, one microscope picture and one MRI picture. All results are with run length, but without consecutive statistical or dictionary compression.

7. FUTURE WORK

Here presented technique can provide comparable quality with simpler framework then wavelet-based techniques. However, there is number of questions which are not answered here. Most of them are concerned about giving the best possible subjective quality for a given CR.

- Proportion of bit rate given to each scale
- When to code some part of residual
- Additional coding on block boundaries
- Deblocking and deringing filters
- Fixed point implementation
- Embedded and progressive ordering
- Supplemental coding with prime numbers block sizes

Solution to number of those questions may lie in concept of profiles. For each respective type of images (photographs, medical, astronomy, industrial, etc.) and CR there should be at least one profile. Then it can be used as near-optimal solution or as better starting point for total optimization in case it is required. Finally, this schema could be extended for volumetric data coding.

8. CONCLUSION

It is shown clearly that DCT-based encoder can compete with wavelet-based encoders, from low to very high CRs, unlike conclusion from [5]. Considering easy implementation, good compression, high quality and other features, this schema could be very interesting in number of areas. Its use is attractive especially in medicine where large images are frequent and where additional efforts can be rewarded in terms of faster display, shorter transmission time and better fidelity. This schema can be good as compression part in medical digital picture archiving and communications system (PACS). Fact that this encoder is far from maturity, while competitive wavelets encoders (JPEG2000 compliant) are fully finished products should be considered also.

<table>
<thead>
<tr>
<th>Size</th>
<th>Ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>Y0(0)-Y0(4095)</td>
</tr>
<tr>
<td>32</td>
<td>[ Y0(0) – Y3(0) ] – [ Y0(1023) – Y3(1023)]</td>
</tr>
<tr>
<td>16</td>
<td>[ Y0(0) – Y15(0) ] – [ Y0(255) – Y15(255)]</td>
</tr>
<tr>
<td>8</td>
<td>[ Y0(0) – Y63(0) ] – [ Y0(63) – Y63(63)]</td>
</tr>
<tr>
<td>4</td>
<td>[ Y0(0) – Y255(0) ] – [ Y0(15) – Y255(15)]</td>
</tr>
</tbody>
</table>

9. References


MULTI-SCALE DCT IMAGE COMPRESSION

Radivoje Vasiljević
Mentor: Prof. Branimir Reljin

From left to right: original, image compressed with JPEG2000 and image compressed with proposed schema at same CR

Figure 1: Lena (at 1:32 – 0.25 bpp) PSNRs are 37.77 for JPEG2000 and 37.56 for proposed schema. Subjective quality is almost identical, as suggested by PSNR.

Figure 2: Microscope image (at 1:128 – 0.0625 bpp) PSNRs are 40.10 for JPEG2000 and 39.72 for proposed schema. JPEG2000 introduced blurring and loss of small details, like bright spot surrounded with larger grey spot at lower left corner. Object right from centre is degraded with edges showing blurred stair-like effect. Finally, content of rhomb like object above centre is very blurred.
Figure 3: Mandrill (at 1:128 – 0.0625 bpp) PSNRs are 29.78 for JPEG2000 and 29.56 for proposed schema. JPEG2000 introduced significant blurring and distortion of image; corners with fur are especially distorted and blurred. Proposed schema degrade image also, but fur is much better preserved. This example clearly shows inadequacy of PSNR metric.

Figure 4: MR (at 1:32 – 0.25 bpp) PSNRs are 41.87 for JPEG2000 and 41.20 for proposed schema. Both images are almost identical to original. JPEG2000 introduced slight blurring that can be explained by fact that source image already was high quality classical JPEG. This example shows that quantisation at lower scales needs more research.