

АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ ЧЕЛОВЕКА К ШУМУ КВАНТОВАНИЯ КОЭФФИЦИЕНТОВ БИОРТОГОНАЛЬНОГО (2,2) ЦЕЛОЧИСЛЕННОГО ДИСКРЕТНОГО ВЕЙВЛЕТ – ПРЕОБРАЗОВАНИЯ КОЭНА–ДОБЕШИ-ФОВО

Мишуровский М.Н., Фартуков А.М., Шерешевский Д.И.

«Юник Ай Сиз»

В последние несколько лет наметились определенные тенденции в развитии алгоритмов сжатия изображений. Многие авторы отмечают, что значительное улучшение качества сжатых изображений может быть достигнуто за счет учета особенностей человеческой системы восприятия визуальной информации (ЧСВВИ) в методах сжатия. Основой многих алгоритмов сжатия видео информации являются дискретные вейвлет преобразования. Вейвлет представление сигнала согласуется с изученными особенностями восприятия визуальной информации: частотная и угловая селективность.

Одним из базовых понятий, применяющихся для математического описания моделей ЧСВВИ, является контрастно – частотная характеристика (КЧХ). Модели маскирования и суммирования искажений основываются на КЧХ. Для оценки деградации изображения при сжатии с потерями, необходимо в первую очередь определить КЧХ, которая учитывает особенности как применяемого вейвлет преобразования, так и используемой системы цветопредставления.

В наших разработках используется реализация биортогонального (2,2) целочисленного дискретного вейвлет – преобразования Коэна – Добеши – Фово и система цветопредставления YCbCr.

Представленное исследование основано на работе группы специалистов под руководством Анджо Ватсона [1]; в статье проведены измерения порогов чувствительности человека к шумам квантования коэффициентов биортогонального вейвлет – преобразования 9-7 Коэна – Добеши – Фово. Это явилось основной причиной для проведения исследований. Методика измерений основана на методе QUEST [2].

Психометрическая функция устанавливает зависимость между вероятностью определенной реакции человека, и какой - либо численной характеристикой раздражителя (или стимула). Психометрическая функция определяется следующим образом (Weibull 1951):

$$p(x) = 1 - e^{-\left(\frac{x}{T}\right)^\beta}, \quad (1)$$

где x – численная характеристика раздражителя (мера), β – «наклон» психометрической функции, T – пороговое значение меры стимула, ε – параметр, определяющий вероятность ответа при $x = T$.

Утверждается, что если выразить меру стимула в логарифмическом диапазоне, то для тестов одного вида психометрическая функция будет иметь одинаковый вид (т.е. параметры T , β , ε - const) с точностью до сдвига. Т.е.

$$p(\bar{x}) = \psi(\bar{x}, \bar{T}, \bar{\varepsilon}) \quad (2)$$
$$\psi(\bar{x}, \bar{T}, \bar{\varepsilon}) = 1 - (1 - \gamma) \cdot e^{\frac{\beta}{20} \cdot (\bar{x} - \bar{T} + \bar{\varepsilon})}$$

Черта над переменной обозначает, что значение выражено в логарифмической шкале - $20\log(x)$. Параметр γ отражает вероятность детектирования сигнала с нулевой мерой. Отражает факт того, что при нулевой мере сигнала человек «угадывает» правильную реакцию на стимул. Величина этого параметра зависит от вида отклика: в случае отклика вида ДА/НЕТ, эта величина равна 0. В случае двух альтернативного выбора (2AFC) – 0.5.

В случае испытаний, связанных с реакцией человека на статическую визуальную информацию, в качестве численной характеристики стимула часто выступает контраст изображения. Само понятие контраста в научном мире четко неопределенно. Возможно поэтому, в [1] в качестве меры стимула выступает амплитуда пикового отклонения стимула от нуля. При этом стимул рассматривается не в пространстве вейвлет – коэффициентов, а в пространстве прямых значений пикселей для рассматриваемого цветового канала.

Методика измерений.

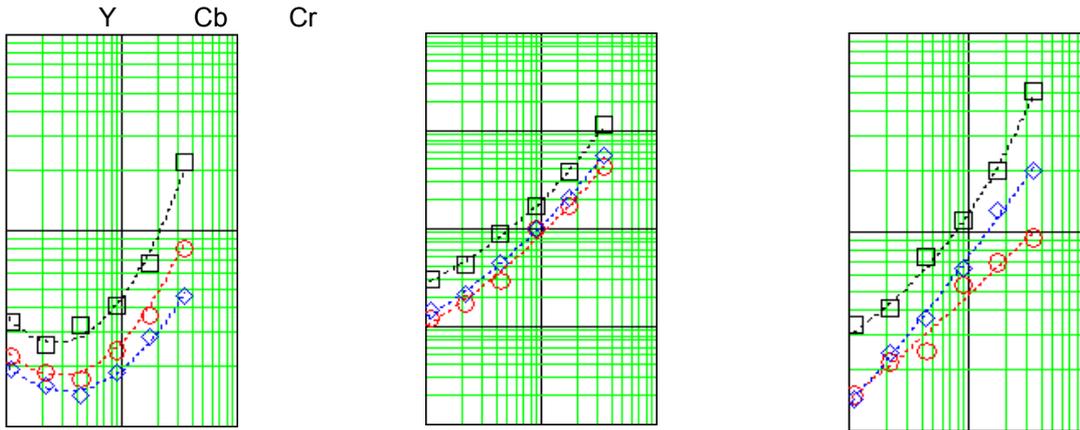
Для проведения измерений порогов чувствительности разработано специализированное программное обеспечение, реализующее метод QUEST и два вида получения отклика: отклик вида ДА/НЕТ и двух альтернативный выбор. По своей сути, первый вид отклика подвержен субъективным факторам и как результат этого - увеличение ошибки измерения. Поэтому все представленные ниже результаты получены с откликом второго типа. Параметр $\beta = 3.5$ (в

литературе указано, что значение этого параметра лежит в диапазоне от 3 до 4). $\gamma = 0.5$; $\varepsilon = -0.91$ Дб, что дает вероятность детектирования порогового сигнала равную 0.75.

Условия наблюдения:

Гамма скорректированный монитор View Sonic P220F, регулировка монитора проведена специализированным ПО фирмы View Sonic. Дистанция наблюдения 60см. Испытуемый – мужчина, 23 лет с нормальным бинокулярным зрением. Условия освещения: постоянное освещение лампами дневного света.

Измерения порогов проведены для высокочастотных подполос для уровней с 1 по 6. Для всех измерений использовалась одинаковая реализация шума, которая инжектировалась в тестируемую подполосу. Результаты эксперимента представлены на графике. По оси X отложена пространственная частота (пиксель/градус). Ось Y – значения порогов.



Математическая модель, описывающая полученные результаты.

Канал светимости (Y): $T(x) = 10^{a1 \cdot (\log_{10} x)^2 + a2 \cdot \log_{10} x + a3}$ (3)

Цветоразностные компоненты: $T(x) = 10^{a1 \cdot e^{a2 \cdot \log_{10} x} + a3}$ (4)

x - пространственная частота в пиксель/градус.

Отличие от модели Ватсона: в его работе цветоразностные компоненты приближались с помощью квадратичной зависимости. Результат такой аппроксимации: спад чувствительности на низких частотах (что противоречит данным экспериментов). Параметры кривых, аппроксимирующих экспериментальные данные:

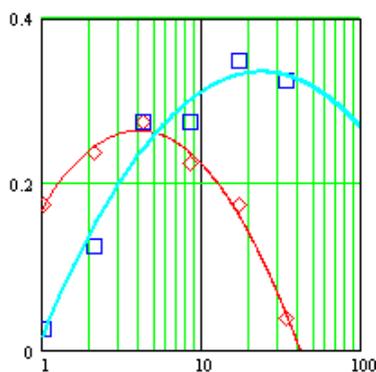
Цвет	HL			LH			HH		
	a1	a2	a3	a1	a2	a3	a1	a2	a3
Y	0.586	-0.554	0.369	0.483	-0.492	0.298	0.714	-0.599	0.547
Cb	1.867	0.411	-1.874	1.805	0.427	-1.73	0.822	0.708	-0.382
Cr	3.443	0.142	-3.285	4.257	0.16	-4.123	0.499	0.802	0.017

Для получения значений порогов в смысле Ватсона, необходимо умножить (3), (4) на 0.7724237 (пиковое значение используемого шума с коэффициентом масштабирования 1).

Обсуждение результатов.

1. Полученные измерения порогов позволяют определить коэффициенты квантования, которые обеспечивают сжатие без «визуальных потерь» с использованием вейвлет преобразования (2,2) Козна – Добеши - Фово при данных условиях наблюдения (максимальная пространственная частота, освещение, гамма = 1). Вычисление коэффициентов квантования осуществляется по формуле (9) из [1] Другой способ вычисления коэффициентов квантования – гипотеза сохранения энергии. Увеличение степени сжатия без потерь при таком способе квантования порядка 50%.

2. Измерение порогов для базисных функций (см. [1], стр. 16-17) и анализ разности между порогами для базисных функций и шумового сигнала показал, что модель Ватсона, объясняющая такое поведение формула (4), [1] корректна не в полной мере. В частности, она не позволяет различать коэффициенты с одинаковыми амплитудами, но разным знаком. Таким образом, пороги для сигнала $\{x_j\}$ и $\{|x_i|\}$ должны быть одинаковы. Результаты исследования утверждают иное.



На представленном графике показано отношение порогов базисной функции к порогам двух сигналов: $\{x_j\}$ и $\{|x_i|\}$, которые были измерены для HL полосы. По оси X отложена пространственная частота в пиксель/градус.

Такой феномен может быть объяснен тем фактом, что амплитудное значение пикового отклонения стимула не является удовлетворительной мерой контраста. Таким образом, встает вопрос о введении адекватного определения термина контраст.

3. Для использования полученных результатов, необходимо определить функцию $P(\{x_j\}, L, O)$, которая определяет вероятность детектирования сигнала $\{x_j\}$ для уровня L подполосы O .

4. Несмотря на то, что в литературе встречается много подтверждений тому, что параметр β инвариантен относительно пространственной частоты и ориентации, имеет смысл определить значения этого параметра для нескольких уровней декомпозиции вейвлет – преобразования. Таких исследований пока не обнаружено.

Благодарности:

Отдельная благодарность Joshua A. Solomon за полезное обсуждение.

Литература:

1. Andrew B. Watson, Gloria Y. Yang, Joshua A. Solomon & John Villasenor "Visibility of wavelet quantization noise".
2. Andrew B. Watson, Dennis G. Pelli "QUEST: A Bayesian adaptive psychometric method"
3. Andrew B. Watson "Probability summation over time", Vision Research, Vol. 19, стр. 515 - 522



ANALYSIS OF HUMAN SENSITIVITY TO QUANTIZATION NOISE OF COEFFICIENTS OF BIORTHOGONAL (2,2) INTEGER DISCRETE WAVELET TRANSFORM COHEN – DAUBECHIES - FEAUVEAU

Mishourovsky M., Fartukov A., Shereshevsky D.

«UniqueICs»

In last years certain tendencies in the field of video coding have been sketched. Several authors point out, that significant improvement of compressed images could be achieved by means of using properties of Human Visual System (HVS). A lot of still image compression algorithms are based on wavelet decomposition. The nature of Wavelet decomposition partially corresponds with significant properties of HVS: directional and frequency selectivity.

One of the main terms, which are applied to describe HVS is Contrast Sensitivity Function (CSF) In order to estimate degradation of image quality during lossy compression, its need to determinate CSF, which take into account features both using certain transform as certain color space.

In our projects we use integer implementation of biorthogonal (2,2) wavelet transform Cohen – Daubechies – Feauveau and YCbCr color space.

This research is inspired by work of Andrew B. Watson et al [1]; They measured human sensitivity to quantization noise of coefficients of biorthogonal 9-7 wavelet transform Cohen – Daubechies – Feauveau. As in [1], we use method QUEST[2] to our measurements.

Observation conditions:

Gamma corrected monitor View Sonic P220f, distance of observation 60cm. Observer – man, 23 years old, normal binocular vision, constant illumination.

Threshold measurements was performed for high frequency bands for levels 1...6. For all measurements the same noise realization was used.

Mathematical model, which describe values of thresholds.

$$\text{Luminance channel(Y): } T(x) = 10^{a1 \cdot (\log_{10} x)^2 + a2 \cdot \log_{10} x + a3} \quad (1)$$

$$\text{Chrominance channels: } T(x) = 10^{a1 \cdot e^{a2 \cdot \log_{10} x} + a3} \quad (2)$$

x - spatial frequency in pixel/degree.

Curve parameters:

	HL			LH			HH		
	a1	a2	a3	a1	a2	a3	a1	a2	a3
Y	0.586	-0.554	0.369	0.483	-0.492	0.298	0.714	-0.599	0.547
Cb	1.867	0.411	-1.874	1.805	0.427	-1.73	0.822	0.708	-0.382
Cr	3.443	0.142	-3.285	4.257	0.16	-4.123	0.499	0.802	0.017

To obtain thresholds as it implies A. Watson, we have to multiple (1), (2) by 0.7724237.

Conclusions.

1. Presented above results allows to perform visual lossless compression (by means of formula (9) in [1]). Another way to obtain quantization coefficients – assumption of “energy preserving” Compression ratio improvement is about 50%.

2. Threshold measurements for single basis function (see [1], pp. 16-17) showed that Watson model of difference between basis function thresholds and noise thresholds are not corrected(see (4), [1]) Especially, it does not allow to distinguish between $\{x_i\}$ and $\{|x_i|\}$ signals.

3. To use mentioned above results its need to define function $P(\{x_j\}, L, O)$, which define probability of detection signal $\{x_j\}$ for level L subband O.

Acknowledgments:

Joshua A. Solomon for helpful discussion.

Reference:

1. Andrew B. Watson, Gloria Y. Yang, Joshua A. Solomon & John Villasenor “Visibility of wavelet quantization noise”.
2. Andrew B. Watson, Dennis G. Pelli “QUEST: A Bayesian adaptive psychometric method”
3. Andrew B. Watson” Probability summation over time”, Vision Research, Vol. 19, pp. 515 - 522