

**Учреждение образования
«Высший государственный колледж связи»
Факультет электросвязи**

Кафедра Программное обеспечение сетей телекоммуникаций

КУРС: ЦИФРОВАЯ ОБРАБОТКА РЕЧИ И ИЗОБРАЖЕНИЯ

для специальности:

1-45 01 03

Сети телекоммуникаций

**Лабораторная работа №3 Критерии оценивания качества
воспроизведения речи и изображений. (8 часов)**

Подготовил:
ст.преподаватель кафедры ПОСТ
Киркоров С.И.

Минск 2010

Лабораторная работа №3 Критерии оценивания качества воспроизведения речи и изображений. (8 часов).

1. Введение

Одним из наиболее мощных программных пакетов обработки мультимедийных данных является MatLab. С его помощью может осуществляться программное моделирование процессов и систем передачи информации. Основные достоинства применения MatLab для программного моделирования состоят в простоте программирования (синтаксис команд интерпретатора MatLab аналогичен синтаксису С-программ), в удобстве отладки программ (отсутствует необходимость в компиляции программы перед ее выполнением, после выполнения программы доступна информация о состояниях всех ее переменных), в простоте и эффективности визуализации результатов выполнения программы (формирование одномерных, двумерных и трехмерных графических объектов), в поддержке большинства стандартных форматов представления и хранения файлов данных, аудио- и видео- информации (неподвижных и подвижных изображений), в наличии развитой библиотеки встроенных функций, реализующих основные операции обработки (информационные преобразования, фильтрация и т.д.) и хорошей документированности встроенных функций (поиск по функциям, описание функций, теоретические сведения по цифровой обработке сигналов, примеры программ).

2.1 Цель лабораторной работы:

Изучение критерий оценивания качества воспроизведения речи и изображений. Закрепление навыков полученных при выполнении лабораторных заданий №1 и №2.

2.2 Критерии оценивания качества воспроизведения речи и изображений

В данном разделе описываются некоторые критерии сравнения изображений. Везде считается, что изображение задается таблицей чисел состоящей из M строк и N столбцов. Каждое число в данной таблице описывает один пиксел, который представляется K битами. Во всех рассматриваемых здесь критериях сравнения степень близости изображений определяется числом, которое некоторым образом вычисляется по данным изображениям. Пусть первое изображение описывается таблицей чисел $\mathbf{X} = \{X_{ij} : 1 \leq i \leq M, 1 \leq j \leq N\}$, а второе –

$\mathbf{Y} = \{Y_{ij} : 1 \leq i \leq M, 1 \leq j \leq N\}$. Для расчета оценок отличия изображений можно использовать следующие выражения:

1. Пиковое отношение сигнал/шум PSNR

$$PSNR = 10 \log_{10} \frac{(2^K - 1)^2}{MSE},$$

где $MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - Y_{ij})^2$ – среднеквадратичная ошибка, K – битовая глубина

цвета, т.е. количество бит требуемое для представления одного пиксела. Обычно вычисляется в децибелах [dB]. Данная мера отличий изображений является классической и, в некотором смысле, эталонной. Во всех работах по сравнению изображений любой нововведенный критерий сравнивается с PSNR. Считается, что если значения PSNR больше чем 37 [dB], то различие изображений практически незаметно, а если меньше 20 [dB], то различие изображений является существенным и весьма значительным. Вместо PSNR иногда используются похожие выражения. Например, среднеквадратичная ошибка MSE, которая записана выше, максимальная среднеквадратичная погрешность PMSE,

$$PMSE = \frac{MSE}{\max |X_{ij}|^2} = \frac{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - Y_{ij})^2}{MN \max |X_{ij}|^2},$$

нормированная среднеквадратичная погрешность NMSE

$$NMSE = \frac{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - Y_{ij})^2}{\sum_{i=1}^M \sum_{j=1}^N (X_{ij})^2},$$

отношение сигнал/шум SNR

$$SNR = 10 \log_{10} \frac{1}{NMSE} = 10 \log_{10} \frac{\sum_{i=1}^M \sum_{j=1}^N (X_{ij})^2}{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - Y_{ij})^2}.$$

Все эти меры отличия изображений дают результаты более или менее похожие с PSNR. Поэтому их рассмотрение не является целесообразным, при условии, что используется PSNR.

2. Максимальная абсолютная ошибка MAE

$$MAE = \max |X_{ij} - Y_{ij}|.$$

Также рассматривается, как классическая мера отличия изображений. Если изображения идентичны, то MAE = 0. Данная метрика хорошо отмечает даже самые малые различия двух изображений, что, с другой стороны, является ее недостатком. Поэтому она не так часто, как PSNR, используется в литературе по оценке отличия изображений.

3. Индекс структурной похожести (SSIM)

$$SSIM = \frac{(2\bar{X}\bar{Y} + C_1)(2\sigma_{XY} + C_2)}{(\bar{X}^2 + \bar{Y}^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)},$$

где $\bar{X} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N X_{ij}$ – среднее значение первого изображения, $\bar{Y} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N Y_{ij}$ –

среднее значение второго изображения, $\sigma_X^2 = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X})^2$ и

$\sigma_Y^2 = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (Y_{ij} - \bar{Y})^2$ – выборочные дисперсии первого и второго изображений

соответственно, $\sigma_{XY} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X})(Y_{ij} - \bar{Y})$ – выборочный коэффициент

корреляции. Неотрицательные константы C_1 и C_2 выбираются, вообще говоря,

произвольно. Данный индекс впервые был введен в работе Z. Wong, A.C. Bovik, H.R.

Sheikh, E.P. Simoncelli *Image quality assessment: from error visibility to structural similarity*

IEEE Transaction on Image Processing, V. 13, No. 3, 2004, 600–612. Эта формула является обобщением индексов SIM и UQI, которые были предложены этими же авторами до этого.

В этой же статье авторами предлагаются следующие значения $C_1 = (0.01L)^2$ и

$C_2 = (0.03L)^2$, где $L = 2^K - 1$, K – битовая глубина цвета, т.е. количество бит требуемое

для представления одного пикселя. Множители перед L в константах C_1 и C_2 следует

выбирать достаточно малыми. SSIM, наряду с его предшественниками SIM и UQI, являются одними из модных в настоящее время эмпирических мер сравнения изображений.

Индекс структурной схожести может принимать значения от -1 до 1. Причем SSIM = 1,

если изображения совпадают. Дополнительную информацию об индексе структурной схожести и ее аналогах можно найти в работах А.С. Bovik, Z. Wong и других (см.

цитируемые работы в указанной выше статье).

4. Приведенная нормализованная мера совместной информации NMIM

$$NMIM = 2 - \frac{H(\mathbf{X}) + H(\mathbf{Y})}{H(\mathbf{X}, \mathbf{Y})},$$

где $H(\mathbf{X}) = -\int_{x \in \mathbf{X}} p_X(x) \log_2 p_X(x) dx$, $H(\mathbf{Y}) = -\int_{y \in \mathbf{Y}} p_Y(y) \log_2 p_Y(y) dy$, –

энтропии соответствующих изображений,

$H(\mathbf{X}, \mathbf{Y}) = -\int_{x \in \mathbf{X}, y \in \mathbf{Y}} p_{XY}(x, y) \log_2 p_{XY}(x, y) dx dy$ – совместная энтропия. Здесь

$p_X(x)$, $p_Y(y)$, $p_{XY}(x, y)$ – соответствующие плотности распределения (или относительные частоты).

В теории информации совместной информацией, содержащейся в двух изображениях, называется число, которое вычисляется по следующей формуле

$MI(\mathbf{X}, \mathbf{Y}) = H(\mathbf{X}) + H(\mathbf{Y}) - H(\mathbf{X}, \mathbf{Y})$. При помощи совместной информации

выражение для NMIM можно записать в таком виде

$$NMIM = 1 - \frac{MI(\mathbf{X}, \mathbf{Y})}{H(\mathbf{X}, \mathbf{Y})}.$$

При такой нормализации NMIM может принимать значения от 0 до 1. Причем NMIM = 0, если изображения совпадают и NMIM = 1, если изображения независимы. Такая метрика впервые вводится в работе Zhang, J., Rangarajan, A.: *Affine image registration using a new information metric*. IEEE Computer Vision and Pattern Recognition (CVPR). Vol. 1. (2004) 848–855. В следующей работе Zhang, J., Rangarajan, A.: *Multimodality image registration using an extensible information metric and high dimensional histogramming* данная метрика применяется для обработки медицинских изображений. Там же приводятся ссылки на работы, в которых используется совместная информация MI при рассмотрении аналогичных вопросов.

Для дискретных изображений NMIM считается следующим образом. Обозначим через

$n_X(a) = \#\{X_{ij} = a\}$ и $n_Y(a) = \#\{Y_{ij} = a\}$ количество пикселей со значением a в изображениях

X и Y соответственно, а через $n_{XY}(a, b) = \#\{X_{ij} = a, Y_{ij} = b\}$ количество пар пикселей $(X_{ij}; Y_{ij})$ в

изображениях X и Y таких, что $X_{ij} = a$ и $Y_{ij} = b$. Тогда $H(\mathbf{X}) = -\sum_{a \in X} \frac{n_X(a)}{NM} \log_2 \frac{n_X(a)}{NM}$ и

$H(\mathbf{Y}) = -\sum_{a \in Y} \frac{n_Y(a)}{NM} \log_2 \frac{n_Y(a)}{NM}$, где суммирование проводится по всем значениям, которые

могут принимать пиксели из соответствующих изображений. Для совместной энтропии имеем

$H(\mathbf{X}, \mathbf{Y}) = -\sum_{a \in X} \sum_{b \in Y} \frac{n_{XY}(a,b)}{NM} \log_2 \frac{n_{XY}(a,b)}{NM}$, здесь суммирование проводится по всем парам

значений, которые могут принимать пиксели из соответствующих изображений.

5. Степень сжатия изображений bpp (bits per pixel).

$bpp = \frac{B}{NM}$, где B — общее количество битов в сжатом изображении, N, M — размеры

изображения по вертикали и горизонтали в пикселях. Данная характеристика является общепринятой для оценки степени сжатия изображений. Так, например, если изображение не

подвергалось сжатию и для хранения одного пикселя использовалось 8 бит, то $bpp = \frac{8NM}{NM} = 8$. Для

8-ми битных изображений при сжатии от 10 до 20 раз bpp меняется от 0.8 до 0.4 соответственно.

2.3 Рекомендации по вычислению выбранных метрик оценки качества изображений

При вычислении метрик MAE, PSNR, SSIM, NMIM, по предлагаемым формулам в случае очень больших изображений, может возникнуть переполнение регистров. Например, при вычислении MSE по формуле $\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (f_1(i,j) - f_2(i,j))^2$ для двух 8-битных квадратных изображений изображений f_1 и f_2 в худшем случае может быть $|f_1(i,j) - f_2(i,j)| = 255$ для всех i, j .

Отсюда $\sum_{i=1}^N \sum_{j=1}^N (f_1(i,j) - f_2(i,j))^2 = 65025N^2$. При использовании плавающей точки типа float, когда на число отводится 32 бита, уже при $N > 5800$ относительная погрешность превосходит 50%. При использовании фиксированной точки неприемлемые результаты получаются при еще меньших величинах N . Есть несколько способов избежать ошибок переполнения.

Способ 1. Для вычислений сумм вида $S_n = \frac{1}{n} \sum_{i=1}^n a_i$ можно воспользоваться рекуррентными

формулами $S_{k+1} = \frac{kS_k + a_{k+1}}{k+1}$, $S_0 = 0$ или $S_{k+1} = S_k + \frac{a_{k+1} - S_k}{k+1}$, $S_0 = 0$. В этом случае при

использовании типа float переполнение не происходит при $n < 2^{30}$. Недостатком этого метода является то, что существенно увеличивается время работы алгоритмы вычисления метрики, т.к. для каждого пикселя нужно выполнять деление типа float.

Способ 2. Допустим, необходимо вычислить $S_n = \frac{1}{n} \sum_{i=0}^{n-1} a_i$. Если n можно представить в виде

$n = MN$, а именно так и есть в случае вычисления предлагаемых метрик (M, N — высота и ширина изображения), то можно воспользоваться следующим алгоритмом:

1. Присвоить $S := 0$;

2. Цикл: для каждого i от 0 до $M - 1$ выполнить:

$$b := \frac{1}{N} \sum_{j=0}^{N-1} a_{ni+j}, \text{ где сумма вычисляется обычным суммированием,}$$

$$S := S + \frac{b - S}{i + 1}, \text{ завершение цикла.}$$

3. Вывести S .

Выбор разбиения $n = MN$ должен осуществляться таким образом, чтобы при вычислении b не возникало переполнения. При естественном разбиении 8-ми битного изображения по строкам и столбцам, это условие выполняется для изображений с $N < 2^{15}$. Способ 2 работает существенно быстрее, т.к. количество делений сокращается в N раз.

При вычислении метрик, в которых участвуют выражения вида $\left(\sum_{i=1}^n a_i \right) / \left(\sum_{i=1}^n b_i \right)$ во

избежание ошибок переполнения при суммировании большого количества слагаемых, рекомендуется вычислять такие выражения по правилу:

$$\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} = \frac{\frac{\sum_{i=1}^n a_i}{n}}{\frac{\sum_{i=1}^n b_i}{n}},$$

где средние арифметические в правой части вычисляются, скажем, способом 2.

3. ЛАБОРАТОРНОЕ ЗАДАНИЕ

3.1. Изучите теоретическую часть.

3.2. Используйте опыт и навыки полученные в при выполнении заданий способы лабораторных работ №1, №2 в MATLAB с мультимедийными данными (Приложение лабораторному практикуму): Example_1_1, Example_1_2, Example_1_3, Example_2_1, Example_2_2, Example_2_3, Example_2_4, Example_3_1, Example_3_2, Example_3_3, Example_3_4, а также программу реализующую алгоритм квантования и используйте набор тестовых изображений.

3.3 Создайте квантованное изображение в соответствии со своим вариантом (вариант изображения берутся по последнему номеру в списке журнала группы 1_1-1_6, 2_1-2_6 и т.д.).

3.4. Вычислите характеристики по критериям в соответствии с разделом 2.2 сравнивая исходное и квантованное изображения.

3.5. Оформите отчет и сделайте выводы.

3.6. Факультативно.

3.6.1 Выполните п.3.3 заменив алгоритм на JPEG и JPEG2000. Выполните сжатие и восстановление изображения (BMP=>JPEG=>BMP; BMP=>JPEG2000=>BMP). Оцените потери в изображении по сравнению с исходным изображением. Постройте графики и таблицы для сравнения с разными степенью сжатия (5,10,20,50,100). Сравните алгоритмы сжатия JPEG и JPEG2000. Постройте графики и таблицы для сравнения. Сделайте выводы.

4. СОДЕРЖАНИЕ ОТЧЕТА

4.1 Решение задач задания.

4.2. Результаты выполнения работы.

4.3. Анализ результатов и выводы.