

На основе использования компонентов указанной интеллектуальной системы студенты изучают:

- принцип автоматического перемещения лазерного источника по двум координатам;
- механизм воздействия мощного лазерного излучения на поверхность твердого тела;
- схемотехнические решения микропроцессорного модуля управления электроприводом механизма перемещения;
- особенности программирования микропроцессоров для систем автоматического управления;
- зависимость глубины гравирования от параметров лазерного излучения и скорости излучателя.

УДК 004.032.26

JPEG-КОМПРЕССИЯ ДАННЫХ ВИДЕО СЕНСОРОВ

студент гр. 714301 Забелло К.В.

Научный руководитель – к.т.н. Ролич О.Ч.

Белорусский государственный университет информатики
и радиоэлектроники
Минск, Беларусь

Бесконтактный контроль и диагностика состояния промышленного оборудования в масштабе реального времени требует принятия своевременных и неотложных действий в случаях появления технологических дефектов. Применение узлов технического зрения в системах неразрушающего бесконтактного контроля обуславливает передачу больших объёмов данных от видео сенсоров, что, в свою очередь, требует операции промежуточного сжатия видео кадров в целях экономии трафика и наискорейшего обнаружения дефектов в процессе работы оборудования без его полной остановки. Motion JPEG как наиболее распространенный стандарт передачи видео данных в технических системах базируется на JPEG-компрессии кадров видео сенсоров.

JPEG – графический формат, который хранит данные изображения в сжатой с потерями форме в виде квантованных частотных коэффициентов [1].

Процесс JPEG-сжатия изображения представлен на рисунке 1.

Большинство форматов изображений использует аддитивную модель RGB, основанную на прибавлении к черному красного, зеленого и синего составляющих. Алгоритм JPEG начинается с перевода цветов изображения из системы RGB в перцепционную модель YCbCr:

$$Y = 0 + 0.299 \cdot R + 0.578 \cdot G + 0.114 \cdot B,$$

$$Cb = 128 - 0.168736 \cdot R - 0.331264 \cdot G + 0.5 \cdot B,$$

$$Cr = 128 + 0.5 \cdot R - 0.418688 \cdot G - 0.081312 \cdot B,$$

где компонент Y – яркостная составляющая – является основным. Данный факт обосновывается биологией глаза человека и физикой света: человеческий глаз реагирует на изменения яркости в значительно большей степени, чем на изменения цветовых составляющих. Поэтому, определенной частью данных компонент Cb и Cr можно пренебречь без существенной потери качества результирующего изображения.



Рис.1. Этапы сжатия JPEG.

Непосредственное пренебрежение цветоразностными компонентами осуществляется в процессе субдискретизации. При использовании субдискретизации 4:2:2 на четыре значения яркостной составляющей приходится два значения цветной компоненты. Данный этап пропускается в случае, когда качество изображения на выходе должно быть высоким, с параметрами субдискретизация 4:4:4.

Восьмибитовые беззнаковые величины (Y,Cb,Cr) в изображении центрируются, т.е. преобразовываются в восьмибитовые знаковые путем вычитания 128. Таким образом, величины, ранее принадлежащие диапазону от 0 до 255, центрируются около нуля.

Изображение делится на блоки 8x8 пикселей, и для каждого блока 8x8 применяется DCT-преобразование:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N},$$

где $0 \leq p \leq M - 1$ и $0 \leq q \leq N - 1$;

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0, \\ \sqrt{\frac{2}{M}}, & p > 0; \end{cases} \quad \alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0, \\ \sqrt{\frac{2}{N}}, & q > 0; \end{cases}$$

Дискретное косинусное преобразование (DCT) – это разновидность преобразования Фурье. Графическое изображение DCT рассматривается как совокупность пространственных волн. Дискретное косинусное преобразование позволяет переходить от пространственного представления изображения к ее спектральному образу и обратно. Воздействуя на спектральное представление изображения, состоящее из гармоник, путем пренебрежения наименее значимых из них, осуществляется балансировка между качеством воспроизведения и степенью сжатия.

Таким образом, DCT-преобразование работает с пространством частот изменения яркости и оттенка. Эти частоты тесно связаны с уровнем детализации изображения. Высокие частоты соответствуют высокому уровню детализации.

После выполнения DCT-преобразования над блоком 8x8 формируется новый блок 8x8, который сканируется по зигзагу.

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

Рис.2. Чтение матрицы по зигзагу.

Смысл зигзагообразного вектора заключается в том, что коэффициенты DCT просматриваются в порядке повышения пространственных частот. Первая величина в векторе соответствует самой низкой частоте в изображении – она обозначается термином DC. Остальная часть коэффициентов DCT обозначается AC.

В полученном на предыдущем этапе отсортированном векторе каждая из 64 величин квантуется, т.е. каждая величина делится на число, определенное в таблице квантования, затем округляется до ближайшего целого:

для ($i = 0; i \leq 63; i++$)

вектор[i] = (округлить) (вектор[i] / таблица_квантования[i] + 0.5)

Ниже приведен пример таблицы квантования для яркости (Y):

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Процесс квантования играет ключевую роль в JPEG сжатии и основывается на "психовизуальном эффекте". Глаз более чувствителен к низким частотам, поэтому на этапе квантования происходит деление элементов матрицы (Y, Cb или Cr), которые находятся ближе к правому нижнему углу (высокие частоты), на соответствующие значения в таблице квантования (большие значения), что приводит к удалению высоких частот.

RLE (Run-Length Encoding – кодирование по длинам серий) кодирует последовательность нулей в полученном квантованном векторе, причем первый коэффициент в векторе не кодируется. Поэтому, для каждой величины, отличающейся от 0, кодируется количество последовательных нулей, предшествующих данной величине, и затем добавляется непосредственная величина.

Вместо хранения фактической величины, стандарт JPEG хранит минимальный размер в битах, необходимый для представления величины (категория величины), а также битно кодированное представление данной величины: Конечный шаг кодировки состоит в кодировании байта по Хаффману, и затем записи результата b битового представления числа в файл JPG.

При сжатии методом JPEG потеря информации происходит на этапе квантования, когда коэффициенты округляются до целых чисел в пределах -2048 ... 2047.

Чем больше значения в матрице квантования, тем больше отбрасывается информации из изображения и тем более плотно оно сжимается. Компромисс состоит в

том, что более высокие значения квантования приводят к худшему качеству изображения.

Литература

1. Ватолин, Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ-МИФИ, 2002. - 384 с.

УДК 303.723

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ВЕКТОРНЫХ ИЗОБРАЖЕНИЙ

студент гр. 714301 Кастюкевич Д.В.

Научный руководитель – к.т.н. Ролич О. Ч.

Белорусский государственный университет информатики
и радиоэлектроники
Минск, Беларусь

Системы распознавания образов всё в большей степени внедряются в сферу неразрушающего контроля и диагностики машинных агрегатов [1]. На изображениях, получаемых от видео сенсоров, зачастую требуется измерять параметры формы объектов, которые предварительно необходимо распознать и выделить из общего плана. Считывание автомобильных номеров, надписей на контейнерах и железнодорожных вагонах, анализ текстовых документов, идентификация символьной информации на печатных платах и электронных компонентах также упирается в автоматизированное распознавание образов [1, 2].

Одним из операторов распознавания объекта на изображении является его сравнение с эталоном, что на практике, как правило, реализуется посредством оператора корреляции. С помощью него предварительно векторизованное исходное (т.е. полученное от видео сенсора) изображение сравнивается с образом эталона [3].

Векторное изображение – это изображение, распределение яркости и/или форма контуров в котором заданы математически. Задача векторизации исходного изображения включает в себя следующие этапы:

- выделение области расположения образа целевого объекта;
- шумоподавление и бинаризация принятого образа;
- сегментация и описание полученного образа целевого объекта математическими зависимостями.

Оптимальный алгоритм сравнения, обеспечивающий наибольшую вероятность правильного различения, заключается в анализе корреляционной функции между образом распознаваемого целевого объекта и набором эталонов.

Дискретная двумерная корреляционная функция представляется следующей конечной суммой:

$$R_{k,l} = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} X_{i+k,j+l} E_{i,j},$$

где X – векторизованное исходное изображение целевого объекта размером $M \times N$; E – векторное изображение эталона, $m \times n$ – размеры матрицы эталона E ; k, l - индексы: $k = 0, \dots, (M - 1), l = 0, \dots, (N - 1)$.

База эталонных образов объектов формируется векторным способом. Их форма задается в виде следующих математических уравнений в декартовой или в полярной системе координат.