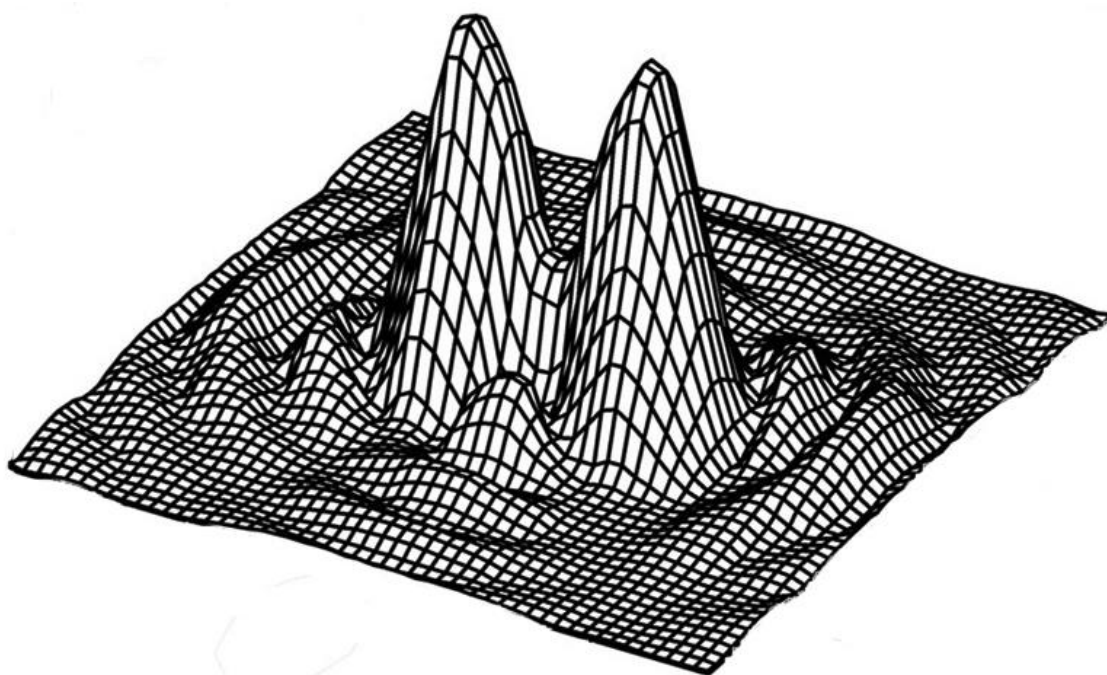


М.А. Волкова, В.Р. Луцив

МЕТОДЫ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Учебно-методическое пособие по
лабораторному практикуму



Санкт-Петербург
2016

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

М.А. Волкова, В.Р. Луцев

МЕТОДЫ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Учебно-методическое пособие по лабораторному
практикуму



Санкт-Петербург

2016

Волкова М.А., Луцив В.Р. Методы обработки и распознавания изображений. Учебно-методическое пособие по лабораторному практикуму. – СПб: Университет ИТМО, 2016. – 40 с.

Пособие содержит краткое описание лабораторного практикума для ознакомления с основами методики обработки и распознавания изображений в системах компьютерного зрения. Даются краткие теоретические сведения о предмете исследования, описывается порядок выполнения лабораторных работ и предъявления их результатов. Пособие адресовано студентам, обучающимся по направлению 09.03.03 (230700) – Прикладная информатика, а также аспирантам и специалистам, интересующимся проблематикой обработки и распознавания изображений

Рекомендовано к печати ученым советом факультета Фотоники и оптоинформатики, протокол № 3 от 07.04. 2016 г.



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО – становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО, 2016

© Волкова Мария Анатольевна, 2016

© Луцив Вадим Ростиславович, 2016

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ.....	4
ЛАБОРАТОРНАЯ РАБОТА №1 Распознавание объекта и определение его декартовых координат на изображении путем вычисления функции взаимной корреляции в пространственно-временной области.....	5
ЛАБОРАТОРНАЯ РАБОТА №2 Распознавание объектов на изображении и определение параметров их пространственного вращения и масштабирования методом корреляции в полярно-логарифмической системе координат.....	10
ЛАБОРАТОРНАЯ РАБОТА №3 Исследование метода измерения параметров аффинного преобразования изображений	18
ЛАБОРАТОРНАЯ РАБОТА №4 Распознавание объектов на изображении методом Фурье-Меллина.....	27
ЛАБОРАТОРНАЯ РАБОТА №5 Распознавание изображений, подвергнутых аффинным преобразованиям.....	34
ПРИЛОЖЕНИЕ	41

ПРЕДИСЛОВИЕ

Пособие содержит описание лабораторного практикума для студентов, начинающих знакомиться основами обработки и распознавания изображений в системах компьютерного зрения. Ознакомление начинается с самых простых методов автоматического анализа изображений, что даст обучающимся возможность самостоятельно убедиться в их ограниченной работоспособности и понять важность применения рассмотренных далее более сложных инструментов обработки изображений. Для каждой лабораторной работы даны краткие теоретические сведения о предмете исследования, описан порядок ее выполнения и предъявления результатов.

Пособие адресовано студентам, обучающимся по направлению 09.03.03 (230700) – Прикладная информатика. В связи с этим целью предлагаемого лабораторного практикума является не только ознакомление с методами обработки изображений в системах компьютерного зрения. В лабораторных работах исследования выполняются на основе самостоятельно разрабатываемых студентами программных моделей, что дополнительно позволяет им совершенствовать навыки программирования. При этом для каждой лабораторной работы в разделе предлагаемых теоретических сведений подробно описывается, как самостоятельно реализовать необходимый метод обработки изображения в виде программной модели, что дает студентам исчерпывающее представление об ее структуре. В то же время чтение изображения из графического файла стандартного формата, отображение результатов обработки на экране и их сохранение в файлах предлагается выполнять с использованием функций свободно распространяемой библиотеки компьютерного зрения Intel OpenCV (пример приведен в приложении). И более того, студентам рекомендуется самостоятельно найти в указанной библиотеке готовые функции, которые могут использоваться в качестве составных блоков создаваемых ими программных моделей. Это позволяет им осваивать самые современные средства компьютерной обработки изображений. Рекомендательный для данного лабораторного практикума язык программирования – C/C++, однако, возможен и самостоятельный выбор студентами среды программирования.

Файлы установочных дистрибутивов для современных версий библиотеки OpenCV можно загрузить с сайта <http://opencv.org/> ее разработчика, на этом же сайте можно найти инструкции по ее установке и использованию. Начинающие пользователи могут воспользоваться первой версией библиотеки – OpenCV 1.0, она проще в установке и использовании. Эту версию и инструкции по ее установке можно найти на сайте <http://robocraft.ru/blog/computervision/265.html>. Полезное пособие по OpenCV 1.0 (Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А.С. Потапов, М.В. Петерсон, В.А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.) можно найти в библиотеке университета ИТМО. Ссылки на другие полезные литературные источники приводятся в описании каждой лабораторной работы.

ЛАБОРАТОРНАЯ РАБОТА №1

Распознавание объекта и определение его декартовых координат на изображении путем вычисления функции взаимной корреляции в пространственно-временной области

1 Теоретическая часть

Одним из простейших и наиболее распространенных путей поиска эталонного образа в исследуемом изображении является вычисление функции их взаимной корреляции $C(x,y)$:

$$C(x,y) = (f \otimes g)_{x,y} = \sum_{k,l} f_{x+k,y+l}^* g_{k,l},$$

где $f_{x,y}$ и $g_{k,l}$ – значения яркости пикселей сопоставляемых исследуемого и эталонного изображений, соответственно, x и y – декартовы координаты пикселя в двумерном поле вычисленной функции взаимной корреляции (номер строки и номер столбца пикселей, соответственно), они же соответствуют величинам взаимного смещения эталонного g и исследуемого f изображений соответственно по осям абсцисс и ординат, $*$ – операция комплексного сопряжения (ее можно опустить, т.к. в данном случае пикселям приписаны действительные значения яркости), k и l – декартовы координаты пикселя в эталонном изображении. Путем поиска максимума построенной таким образом двумерной корреляционной функции можно определить местонахождение искомого эталона (объекта) в исследуемом изображении. Значение корреляционной функции в точке ее максимума характеризует степень сходства сопоставляемых изображений. Графическая иллюстрация взаимного смещения сопоставляемых изображений в процессе вычисления функции их взаимной корреляции приведена на рис. 1.1. Искомое изображение g помещается в левый верхний угол (положение a в строке α) анализируемого изображения f , перемножаются значения совпадающих по положению пикселей сопоставляемых изображений, и произведения, вычисленные для каждого пикселя искомого изображения суммируются. Сформированная сумма равна одному отсчету корреляционной функции, полученному для данного положения искомого изображения на анализируемом. Искомое изображение g смещается по строке в следующую позицию b , и вычисляется следующий отсчет функции корреляции путем аналогичного суммирования произведений совпадающих по координатам пикселей изображений g и f . Последующие отсчеты первой строки двумерной функции взаимной корреляции вычисляются при перемещении изображения g с шагом 1 по первой строке α до последней позиции z этой строки. Затем изображение g смещается на один пиксел вниз, и аналогичным образом формируются «пиксели» второй строки (строка β) корреляционной функции. Аналогично шаг за шагом по вертикали формируются остальные

строки «пикселей» корреляционной функции вплоть до последней ее строки ω .

Сопоставление изображений путем вычисления функции их взаимной корреляции как частный случай применения согласованного фильтра является оптимальным решением при высоких уровнях аддитивной шумовой составляющей, отличающей изображение реального объекта от эталонного. Пример искомого (эталонного) изображения и анализируемого неизвестного изображения, в котором ищется эталонное, приведен на рис. 1.2.а и 1.2.б, соответственно. Пример корреляционного поля этих двух изображений, рассчитанного выше описанным способом, показан на рис. 1.3.а. Отчетливо виден пик корреляционного отклика в позиции, где на анализируемом изображении находится искомый фрагмент. Однако корреляционный отклик разрушается, если взаимное геометрическое преобразование сопоставляемых изображений описывается более сложными трансформациями системы координат, чем простые взаимные линейные смещения вдоль декартовых координатных осей. Выполняя настоящую работу, студенты должны научиться использовать функцию взаимной корреляции для нахождения эталонного объекта в исследуемом изображении и определить применимость этого метода, если сопоставляемые изображения имеют также взаимные масштабные преобразования и вращения.

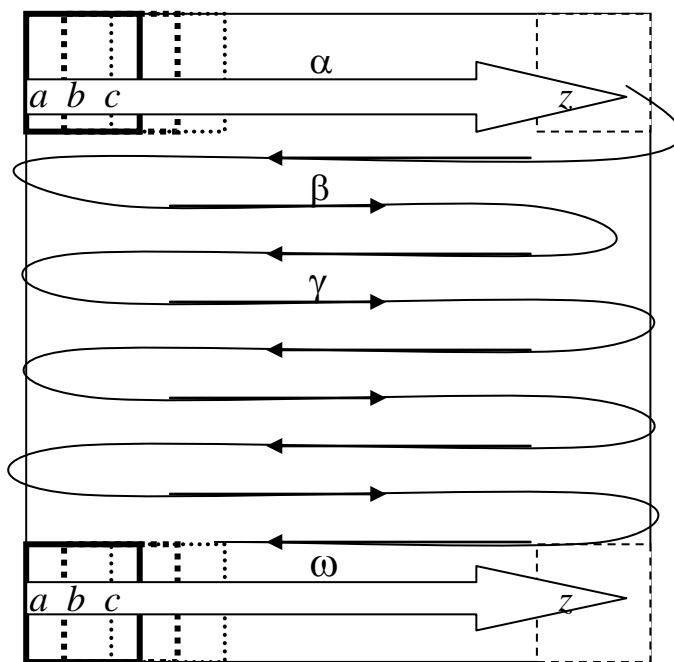


Рисунок 1.1 – Порядок формирования строк и столбцов массива взаимной корреляционной функции двух сопоставляемых изображений

2 Порядок выполнения лабораторной работы

2.1. Получить у преподавателя файл подлежащего исследованию «большого» изображения и файлы пары изображений искомых (эталонных) объектов, которые должны быть обнаружены в исследуемом «большом»

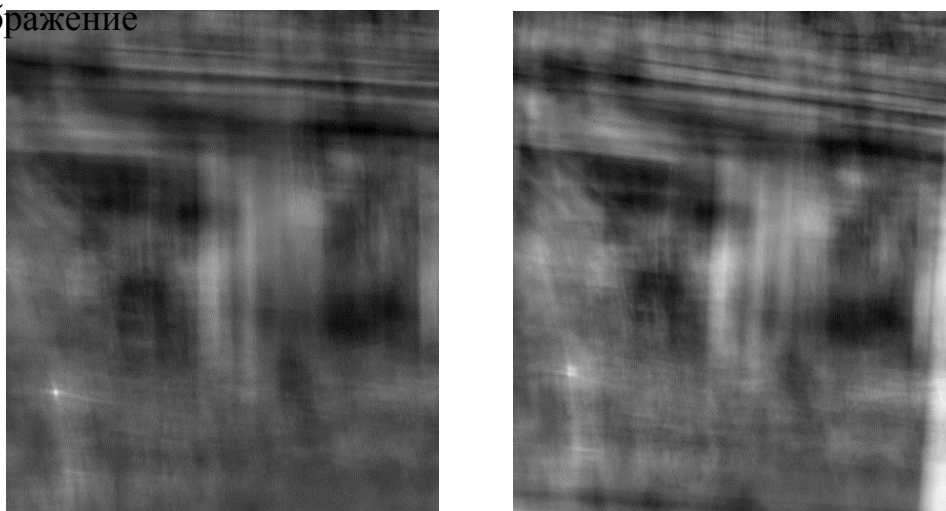
изображении. Перевести полученные изображения, из файлов стандартного графического формата в двумерные массивы пикселей, пригодные для компьютерной обработки. Студенты могут воспользоваться для чтения пикселей из полученных файлов изображений серого тона стандартного графического формата в двумерные массивы, для визуализации массивов обработанных пикселей и для записи массивов в файлы изображений стандартного формата соответствующими функциями библиотеки Intel Open CV; в приложении приведен один из возможных вариантов фрагмента программы, читающей массив пикселей из файла и визуализирующий обработанный массив, см. комментарии, приведенные в приложении.



а

б

Рисунок 1.2 – Пример поиска заданного фрагмента в анализируемом неизвестном изображении: а – искомый фрагмент, б – анализируемое изображение



а

б

Рисунок 1.3 – Корреляционная функция изображений, представленных на рис. 1.2: а) не искаженный корреляционный отклик, б) деградация отклика при взаимном повороте сопоставляемых изображений на 6°

- 2.2.** Вычислить взаимные корреляционные функции исследуемого изображения с каждым из искомых (эталонных) как функции их взаимных смещений вдоль декартовых координатных осей. Определить по виду вычисленных корреляционных функций (корреляционных полей), какой из искомых объектов присутствует в исходном изображении и какова его позиция в нем. Оценить время вычисления корреляционного поля как функцию линейных размеров сопоставляемых изображений. Вычисление функций взаимной корреляции производить без применения быстрого преобразования Фурье (БПФ). Визуализацию рассчитанного корреляционного поля производить путем вывода его в графическое окно на экране терминала, или путем его записи в файл стандартного графического формата.
- 2.3.** Выполнить пункт 2.2 при взаимном вращении исследуемого и искомого изображения (реально найденного в исследуемом при выполнении пункта 2.2), на $\pm 10^\circ$ с шагом 2° (вращение искомого изображения выполнить средствами графических пакетов “Photoshop” или “GIMP”). Оценить, как меняются свойства корреляционного отклика при взаимном вращении сопоставляемых изображений.
- 2.4.** Выполнить пункт 2.2 при взаимном масштабировании сопоставляемых изображений (исследуемого и найденного в нем искомого при выполнении пункта 2.2) с коэффициентом масштабирования, меняющимся в диапазоне от 0,9 до 1,1 с шагом 0,025. Оценить, как меняются свойства корреляционного отклика при взаимном масштабировании сопоставляемых изображений.
- 2.5.** Предъявить результаты выполнения лабораторной работы преподавателю.

Примечание: Во избежание появления больших корреляционных откликов в местах большой постоянной яркости анализируемых изображений перед вычислением функции корреляции сопоставляемые изображения необходимо «центрировать по яркости», т.е. рассчитать среднюю яркость по пикселям массива считанного из файла и из каждого пикселя вычесть рассчитанную среднюю яркость изображения. Рекомендуется также увеличить резкость сопоставляемых изображений, используя фильтр “sharpen” пакета Photoshop или аналогичный фильтр пакета GIMP.

3 Вопросы для самопроверки

1. Почему при вычислении функции корреляции изображений можно опускать операцию комплексного сопряжения?
2. Зачем выполняется «центрирование изображения по яркости»?
3. Что характеризует высота пика функции корреляции двух изображений?

Литература

1. Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А. С. Потапов, М. В. Петерсон, В. А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.
2. Bradski G. and Kaehler A. Learning OpenCV. Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media, Inc., 2008. 556 p.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. 1072 с.

ЛАБОРАТОРНАЯ РАБОТА №2

Распознавание объектов на изображении и определение параметров их пространственного вращения и масштабирования методом корреляции в полярно-логарифмической системе координат

1 Теоретическая часть

Одним из простейших и наиболее распространенных путей поиска эталонного образа в исследуемом изображении является вычисление функции их взаимной корреляции $C(x,y)$:

$$C(x,y) = (f \otimes g)_{x,y} = \sum_{k,l} f_{x+k,y+l}^* g_{k,l},$$

где $f_{x,y}$ и $g_{k,l}$ – значения яркости пикселей сопоставляемых исследуемого и эталонного изображений, соответственно, x и y – декартовы координаты пикселя в исследуемом изображении или в двумерном поле вычисленной функции взаимной корреляции (номер строки и номер столбца пикселей, соответственно), $*$ – операция комплексного сопряжения (ее можно опустить, т.к. в данном случае пикселям приписаны действительные значения яркости), k и l – декартовы координаты пикселя в эталонном изображении. Путем поиска максимума построенной таким образом двумерной корреляционной функции можно определить местонахождение искомого эталона (объекта) в исследуемом изображении. Значение корреляционной функции в точке ее максимума характеризует степень сходства сопоставляемых изображений.

Сопоставление изображений путем вычисления функции их взаимной корреляции как частный случай применения согласованного фильтра является оптимальным решением при высоких уровнях аддитивной шумовой составляющей, отличающей изображение реального объекта от эталонного. Однако корреляционный отклик разрушается, если взаимное геометрическое преобразование сопоставляемых изображений описывается более сложными трансформациями системы координат, чем простые взаимные линейные смещения вдоль декартовых координатных осей. В то же время для некоторых видов преобразования системы координат изображения эта проблема может быть решена. В этой лабораторной работе рассматривается, каким образом можно применять корреляционную функцию для сопоставления изображений, имеющих взаимные масштабные преобразования и вращения.

Для достижения поставленной цели представим изображение в полярной системе координат. В этом случае координаты пикселя на плоскости описываются длиной R радиуса-вектора \mathbf{R} , соединяющего центр пикселя с началом координат, и направлением радиуса-вектора, задаваемым углом ϕ азимута. Развернем окружности равного радиуса в параллельные горизонтальные прямые. Это дает возможность представить азимут и радиус

полярной системы координат ортогональными осями декартовой системы координат, как это показано на рис. 2.1.

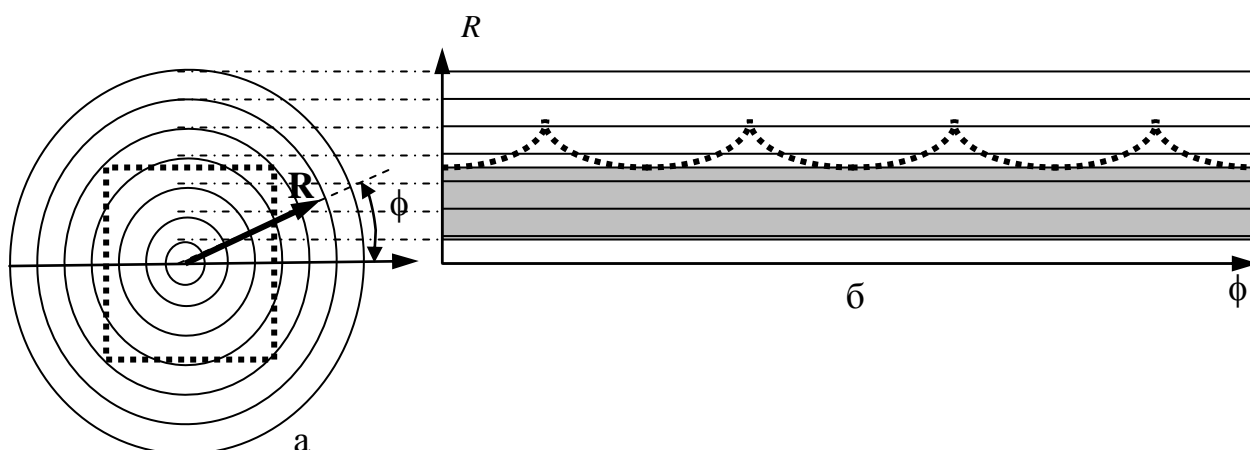


Рисунок 2.1 – Представление азимута и радиуса полярной системы координат с помощью декартовых координатных осей

Представим теперь радиальную координату R на рис. 2.1.б в логарифмическом масштабе [$R' = \ln(R)$] и назовем такую систему координат полярно-логарифмической. Декартовы координаты $\{x, y\}$ пиксела могут быть преобразованы в полярно-логарифмические с помощью следующего преобразования T :

$$\phi = \arctg(y/x); \quad (2.1)$$

$$R' = \ln[(x^2 + y^2)^{1/2}]. \quad (2.2)$$

Тогда масштабирование изображения в исходной декартовой системе координат в M раз приведет к следующему изменению радиальной координаты в полярно-логарифмической системе координат:

$$\begin{aligned} R'(M) &= \ln\{[(Mx)^2 + (My)^2]^{1/2}\} = \\ &= \ln\{M[(x)^2 + (y)^2]^{1/2}\} = \ln(M) + \ln[(x^2 + y^2)^{1/2}]. \end{aligned} \quad (2.3)$$

Таким образом, из выражения (2.3) следует, что масштабирование исходного изображения, представленного в декартовой системе координат, приводит к линейному смещению вдоль оси радиуса его образа, представленного в полярно-логарифмической системе координат, на логарифм коэффициента масштабирования. Из рис. 2.1 также следует, что вращение исходного изображения вокруг начала декартовой системы координат приводит к линейному циклическому смещению вдоль оси азимута его образа, представленного в полярно-логарифмической системе координат. Следовательно, если два изображения, которые мы хотим сопоставить, взаимно масштабированы и повернуты относительно начала декартовой системы координат, но не претерпели линейных смещений в этой системе координат, они могут быть успешно сопоставлены с помощью вычисления их функции взаимной корреляции в полярно-логарифмической системе координат. При этом антилогарифм радиальной координаты пика корреляции

покажет величину взаимного масштабирования, а азимутальная координата пика соответствует величине угла взаимного вращения сопоставляемых изображений. Примеры взаимно повернутых и масштабированных изображений, представленных в декартовой и полярно логарифмической системах координат показаны на рис. 2.2.

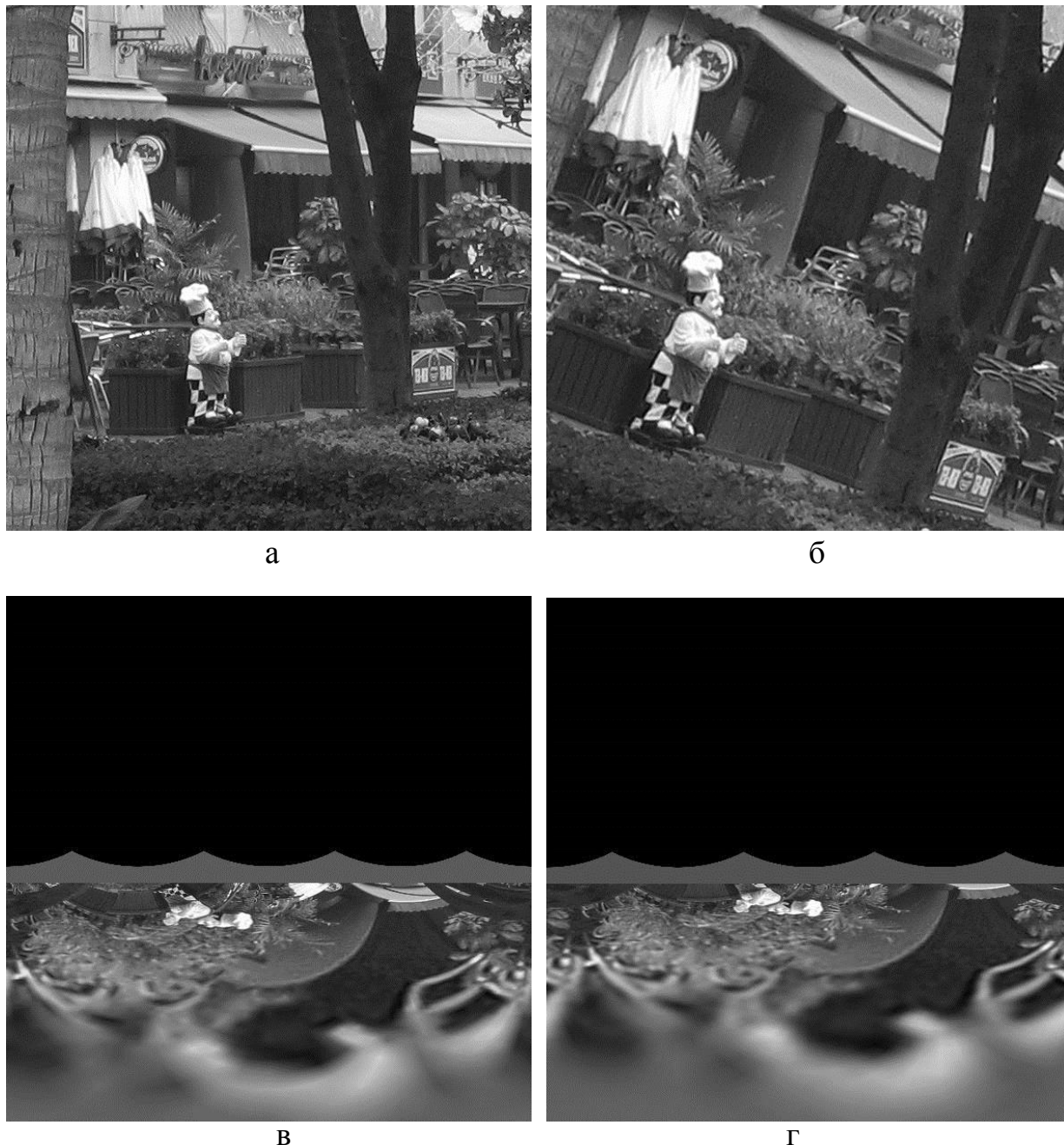


Рисунок 2.2 – Взаимно повернутые и масштабированные изображения (а) и (б), и их представления (в) и (г) в полярно-логарифмической системе координат, взаимно сдвинутые по координатным осям угла и радиуса

Выполняя настоящую работу, студенты должны научиться использовать функцию взаимной корреляции, вычисленную в полярно-логарифмической системе координат для нахождения эталонного объекта в исследуемом изображении и измерения параметров взаимного вращения и масштабирования сопоставляемых изображений.

При выполнении масштабирования и вращения сопоставляемых изображений или при переводе исходных изображений в полярно-логарифмическую систему координат часто приходится вычислять яркость пиксела преобразованного изображения путем интерполяции яркости нескольких пикселей исходного изображения. Аналогичная ситуация возникает при выполнении геометрических преобразований системы координат в настоящей лабораторной работе. Для правильного перевода изображения в полярно-логарифмическую систему координат выполним следующую последовательность операций:

1.1. Пусть исходное (преобразуемое) изображение задано двумерным массивом пикселей $B_{i,j} = B(\mathbf{x})$, где B – яркость пиксела, i – номер строки пикселей в изображении, j – номер пиксела в строке (номер столбца пикселей), $\mathbf{x} = (i, j)^T$ – вектор координат рассматриваемого пиксела.

Внимание!!! Во всех дальнейших вычислениях и преобразованиях предполагается, что номер строки в изображении отсчитывается снизу вверх, в то время как в реальном массиве яркостей пикселей, прочитанном из графического файла, номер строки пикселей нарастает от нуля для самой верхней строки изображения до $N-1$ для самой нижней строки (где N число строк пикселей). Это необходимо учесть в ниже описанных преобразованиях при вычислении ординаты пикселей и направлений угла вращения и угла азимута.

Разместим центральный пиксел преобразуемого изображения в начале декартовой системы координат. Например, если изображение имеет 15 строк ($i = 0 \dots 14$) и 15 столбцов ($j = 0 \dots 14$) пикселей, поместим его пиксел, находящийся в 7-ом столбце 7-й строки, в начало системы координат. Тогда в центрированном таким образом изображении минимальный и максимальный номера строк равны $i_{\min}^c = -7$, $i_{\max}^c = 7$. Минимальный и максимальный номера столбцов будут равны $j_{\min}^c = -7$, $j_{\max}^c = 7$, соответственно. Т.е. для рассматриваемого упрощенного случая, когда число строк и столбцов исходного изображения нечетны, $i^c = i - \frac{N-1}{2}$, $j^c = j - \frac{M-1}{2}$, где N и M – соответственно, число строк и столбцов исходного (преобразуемого) изображения.

1.2. Применим к изображению, центрированному согласно пункту 1.1, преобразование T в полярно-логарифмическую систему координат, описываемое формулами (2.1) и (2.2), и рассчитаем минимальные и максимальные значения радиальной и азимутальной координат, пикселей преобразованного таким образом изображения (минимальные и максимальные номера строк и столбцов пикселей, получаемые в результате преобразования). При этом следует использовать шаг по

угловой координате, равный 1° , что даст азимутальный размер преобразованного изображения (размер вдоль оси абсцисс), равный 360 пикселям (номера столбцов преобразованного изображения меняются от 0 до 359). Радиальный размер изображения, преобразованного в полярно-логарифмическую систему координат (размер вдоль оси ординат), вычисляется как логарифм максимального значения радиальной координаты и увеличивается в 50 раз, чтобы обеспечить достаточно большое количество пикселей вдоль оси радиусов. Затем полученное нецелое число строк округляется в сторону большего по модулю целого значения, это и будет максимальный номер строки массива, в котором будут храниться пиксели преобразованного изображения.

1.3. Вычисляем значения (яркости) каждого пикселя преобразованного изображения (размеры его массива в пикселях были рассчитаны выше согласно пункту 1.2) методом обратного проецирования координат субпикселей. Для этого выполним следующие действия:

1.3.1. Представим каждый пиксел преобразованного изображения как площадку, состоящую из $Q \times Q$ субпикселей ($Q \in \mathbb{N}$), как это показано на рис. 2.3. Пусть Q нечетно ($Q = 2P + 1$, $P \in \mathbb{N}$), например, $Q = 9$, тогда центральному пикселу площадки будут соответствовать собственные целые координаты $\{I, J\}$ пикселя преобразованного изображения, а для любого субпикселя внутри площадки нецелые значения координат вычисляются по формуле

$$\begin{pmatrix} i \\ j \end{pmatrix} = \begin{pmatrix} I + \frac{k}{2P+1} \\ J + \frac{l}{2P+1} \end{pmatrix}, \quad (2.4)$$

где $k = [-P \dots +P]$, $l = [-P \dots +P]$ – целые номера субпикселей по горизонтали и вертикали внутри площадки, соответственно.

1.3.2. Вычислим координаты обратной проекции каждого субпикселя, сформированного согласно пункту 1.3.1, из преобразованного изображения в преобразуемое изображение, центрированное согласно пункту 1.1. Для этого применим к координатам каждого субпикселя, вычисленным согласно формуле (2.4), преобразование T^{-1} , обратное преобразованию в полярно-логарифмическую систему координат, описанному формулами (2.1) и (2.2): $x = \exp(R'/50)\cos(\phi)$, $y = \exp(R'/50)\sin(\phi)$. Вычисленные нецелые координаты обратных проекций субпикселей округляем до ближайших целых значений (соответствующих реальным целым координатам одного из пикселей исходного изображения). Каждому субпикселу с координатами $\{k, l\}$ рассматриваемого пикселя преобразованного изображения присваивается яркость $b_{k,l}$, равная яркости пикселя, которому

соответствует его проекция на исходном изображении, или нулевая яркость, если он спроецировался вне границ исходного изображения.

1.3.3. Яркость B' каждого пиксела преобразованного изображения вычисляется как среднее яркостей $b_{k,l}$ всех соответствующих ему субпикселов:

$$B'_{I,J} = \frac{1}{Q^2} \sum_k \sum_l b_{k,l},$$

где значения яркостей субпикселов $b_{k,l}$ были вычислены методом обратной проекции согласно пункту 1.3.2.

2 Порядок выполнения лабораторной работы

2.1. В данной работе для проведения исследования используется «большое» исходное изображение, рассмотренное в лабораторной работе №1. Оно имеет нечетный размер $2R+1$ ($R \in \mathbb{N}$) по строкам и столбцам пикселов, причем после центрирования изображения согласно пункту 1.1 вертикальные и горизонтальные границы изображения будут иметь координаты $\pm R$, соответственно.

2.2. Центрированное согласно пункту 2.1 «большое» квадратное изображение подвергается масштабированию с коэффициентом $K = 1+0.05 \times N$ и вращению на $(+3 \times N)^\circ$ относительно его центрального пиксела, где N – номер варианта выполнения работы. Масштабирование и вращение можно производить средствами подходящего графического редактора (например, Adobe Photoshop или GIMP) при работе с исходным файлом изображения, представленным в одном из стандартных графических форматов, или написав собственную программу, осуществляющую заданную геометрическую трансформацию. В последнем случае для вычисления яркости пикселов и размеров преобразованного изображения использовать метод обратного проецирования, описанный выше в пунктах 1.1–1.3.

2.3. Представить два изображения, сформированные согласно пунктам 2.1 и 2.2, в полярно-логарифмической системе координат согласно методике, проиллюстрированной рисунками 2.1 и 2.2 и формулами (2.1) и (2.2). При преобразовании в полярно-логарифмическую систему координат использовать метод обратного проецирования, описанный выше в пунктах 1.1–1.3. При этом в полярно-логарифмическую систему координат будут переводиться только пикселы, лежащие внутри круга максимального радиуса R , вписанного в каждое из исходных центрированных квадратных изображений. Для пикселов, чья радиальная координата меньше 10, яркость вычисляться не будет, вместо этого каждому из них

приписывается среднее значение яркости, вычисленное по остальной части исходного изображения.

- 2.4.** Вычислить корреляционное поле двух изображений, преобразованных в полярно-логарифмическую систему координат согласно пунктам 2.1–2.3. При вычислении корреляционного поля используется такой же алгоритм перемещения меньшего изображения по большему, как тот, что был применен в лабораторной работе №1, с той лишь разницей, что взаимное перемещение по угловой координате должно быть выполнено циклически. При этом правая часть малого изображения, которая «выдвигается» за правую границу большого, в тот же момент циклически «вдвинется» в большое изображение с левой его стороны. Таким образом, корреляционное поле будет иметь размер 360 пикселей по угловой координате. Размер корреляционного поля по радиальной координате будет равен увеличенной на 1 разнице радиальных размеров сопоставляемых изображений, представленных в полярно-логарифмической системе координат.
- 2.5.** Проанализировать полученное корреляционное поле. Найти максимум корреляции и по его положению определить величину взаимного вращения и масштабирования сопоставляемых изображений. Вывести изображение корреляционного поля в файл стандартного графического формата или графическое окно для его визуализации.
- 2.6.** Предъявить результаты выполнения лабораторной работы преподавателю.

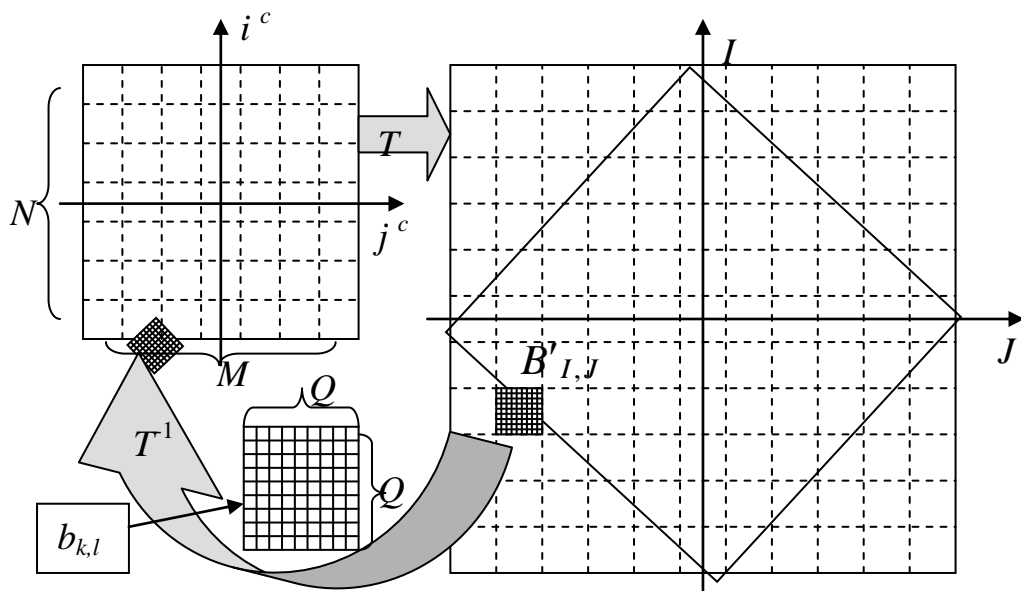


Рисунок 2.3 – Вычисление яркости пиксела в преобразованном изображении субпиксельным методом обратного проецирования (для наглядности вместо преобразования в полярно-логарифмическую систему координат здесь показано преобразование масштабированием и вращением)

3 Вопросы для самопроверки

1. В лабораторной работе №1 наблюдалась деградация корреляционного отклика при взаимном вращении и масштабировании сопоставляемых изображений. Как этот эффект преодолевается в данной работе?
2. Какое геометрическое преобразование используется при вычислении яркость пикселей в полярно-логарифмической системе координат по пикселям исходного?
3. Какой способ интерполяции применен при преобразовании координат?

Литература

1. Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А. С. Потапов, М. В. Петерсон, В. А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.
2. Bradski G. and Kaehler A. Learning OpenCV. Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media, Inc., 2008. 556 p.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. 1072 с.

ЛАБОРАТОРНАЯ РАБОТА №3

Исследование метода измерения параметров аффинного преобразования изображений

1 Теоретическая часть

Для описания зрительных систем высших животных и человека и большинства систем компьютерного зрения могут быть применены схемы, представленные на рис. 3.1.а, рис. 3.1.б и рис. 3.1.в, соответственно. В них точки поверхностей объектов реального мира проецируются с помощью сферической линзы на светочувствительные элементы (палочки и колбочки на внутренней поверхности глаза или элементы светочувствительной матрицы). По законам геометрической оптики такое проецирование осуществляется через точку, соответствующую центру линзы, т.е. проецирование производится непараллельными лучами. В целях упрощения анализа считается, что проецируются точки плоской поверхности объекта (что для большинства объектов реального мира выполняется по крайней мере локально) на плоскую матрицу светочувствительных элементов (что обычно соответствует истине, если мы имеем дело с системами компьютерного зрения, и справедливо для локальных участков сетчатки глаза).

1.1. Если проецируемая поверхность объекта параллельна плоскости, содержащей светочувствительные элементы, как это показано на рис. 3.1.а и рис. 3.1.б, сформированное изображение объекта, подлежащее дальнейшей обработке и распознаванию, геометрически трансформируется преобразованием подобия, включающим в себя линейные сдвиги вдоль декартовых координатных осей, равномерное масштабирование, вращение в плоскости изображения и зеркальное отражение относительно произвольной прямой, лежащей в плоскости изображения и проходящей в ней через начало координат. Масштабирование, вращение и линейные смещения проекции объекта на светочувствительной среде возникают при соответствующих пространственных перемещениях объекта в поле зрения системы относительно его эталонного пространственного положения. Зеркальное отражение, если не рассматривать тривиальное отражение через зеркало, может возникать при наблюдении прозрачного объекта с обратной (задней) стороны, что в реальных ситуациях встречается нечасто.

1.2. Если плоскость, содержащая светочувствительные элементы, становится не параллельна отображаемой поверхности объекта (поворачивается относительно некоторой оси, лежащей в этой плоскости), но объект находится далеко от объектива (линзы), к указанным выше составляющим преобразования подобия добавляется сжатие изображения вдоль прямой, перпендикулярной указанной оси. Полученное более сложное геометрическое преобразование называется аффинным.

Аффинное геометрическое преобразование описывается при этом следующим выражением:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \Leftrightarrow \begin{cases} x' = a_1x + a_2y + a_3 \\ y' = a_4x + a_5y + a_6 \end{cases}, \quad (3.1)$$

где x, y – абсцисса и ордината точки исходного изображения, x', y' – преобразованные координаты этой точки, $a_1 \dots a_6$ – параметры преобразования.

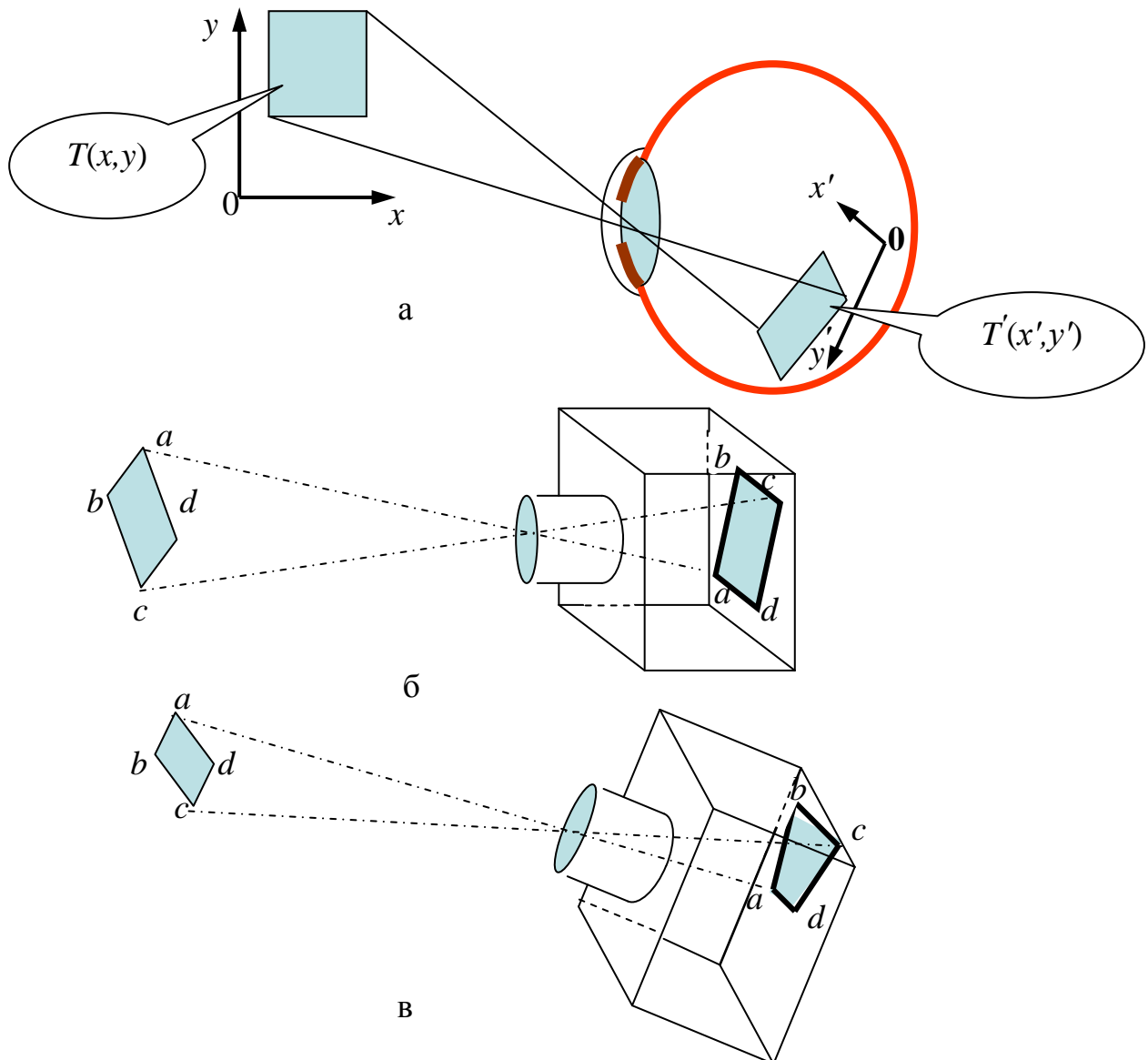


Рисунок 3.1 – Проецирование изображений в искусственных и естественных зрительных системах:

а – проецирование изображения объекта на сетчатку глаза;

б – проецирование в фотокамере при параллельном расположении поверхности объекта и плоскости проецирования;

в – проецирование в фотокамере при непараллельном расположении поверхности объекта и плоскости проецирования.

1.3. Геометрическое преобразование, описанное выше в п. 1.2, еще более усложняется, если отображаемый объект приближается к объективу камеры (хрусталику глаза) на дистанцию, сопоставимую с его фокусным расстоянием (см. рис. 3.1.в). Этот случай соответствует проективному геометрическому преобразованию изображения, описываемому формулами (3.2), включающими два дополнительных параметра a_7 и a_8 . Однако в большинстве практически значимых случаев наблюдаемые объекты находятся достаточно далеко от объектива камеры, при этом значения параметров a_7 и a_8 стремятся к нулю, что приводит формулы (3.2) к виду (3.1).

$$\begin{cases} x' = (a_1x + a_2y + a_3) / (a_7x + a_8y + 1) \\ y' = (a_4x + a_5y + a_6) / (a_7x + a_8y + 1) \end{cases} \quad (3.2)$$

1.4. Как указывалось выше, аффинное преобразование можно представить в виде последовательности более простых преобразований:

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4 \mathbf{A}_5 \mathbf{A}_6, \quad (3.3)$$

где \mathbf{A}_1 соответствует линейному смещению центра тяжести изображения вдоль оси абсцисс, \mathbf{A}_2 соответствует смещению центра тяжести изображения вдоль оси ординат, \mathbf{A}_3 соответствует сжатию вдоль некоторого направления (см. выше в п.1.2), \mathbf{A}_4 соответствует зеркальному отражению, \mathbf{A}_5 соответствует масштабированию, равномерному по двум декартовым координатам, \mathbf{A}_6 соответствует вращению объекта в плоскости изображения относительно начала координат. В общем случае при данном разложении аффинного преобразования порядок его составляющих имеет существенное значение. Однако операции \mathbf{A}_1 и \mathbf{A}_2 ортогональны и могут быть переставлены друг с другом местами. Операции \mathbf{A}_5 и \mathbf{A}_6 также ортогональны и могут быть переставлены друг с другом местами. Операция зеркального отражения \mathbf{A}_4 может выполняться независимо от операции \mathbf{A}_3 – до нее или после нее – а также до или после операции \mathbf{A}_5 (она ортогональна по отношению к операции \mathbf{A}_5). Однако выполнению зеркального отражения перед выполнением операции вращения \mathbf{A}_6 , или после нее будут соответствовать разные значения параметра вращения в разложении (3.3). Пусть для определенности операция зеркального отражения выполняется первой (до операции \mathbf{A}_6).

1.5. Перед началом распознавания изображений (для обеспечения успеха операции распознавания) целесообразно, если это представляется возможным, измерить параметры его аффинного преобразования

относительно его эталонного состояния и компенсировать это преобразование. В соответствии с приведенным в п. 1.4 разложением аффинного преобразования измерение и компенсация составляющих его простых преобразований должны выполняться в следующей последовательности: $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{A}_5, \mathbf{A}_6$. Тогда, если

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4 \mathbf{A}_5 \mathbf{A}_6 \begin{pmatrix} x \\ y \\ 1 \end{pmatrix},$$

в результате компенсации получим

$$\begin{aligned} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} &= \mathbf{A}_6^{-1} \mathbf{A}_5^{-1} \mathbf{A}_4^{-1} \mathbf{A}_3^{-1} \mathbf{A}_2^{-1} \mathbf{A}_1^{-1} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \\ &= \mathbf{A}_6^{-1} \mathbf{A}_5^{-1} \mathbf{A}_4^{-1} \mathbf{A}_3^{-1} \mathbf{A}_2^{-1} \mathbf{A}_1^{-1} \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4 \mathbf{A}_5 \mathbf{A}_6 \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \equiv \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \end{aligned}$$

В приведенных ниже материалах предлагается для исследования метод измерения параметров аффинного преобразования, основанный на сравнении изображения объекта с обобщенными эталонами [1, 2]. Как было сказано в [1, 2] и выше в п. 1.4, эталонным состоянием объекта при таком способе разложения его геометрического преобразования на простые составляющие является такое его пространственное положение, при котором центр тяжести его изображения расположен в начале координат.

1.6. Таким образом, для выполнения операций \mathbf{A}_1^{-1} и \mathbf{A}_2^{-1} , необходимо вычислить абсциссу и ординату (x_c и y_c) центра тяжести изображения объекта как

$$x_c = \frac{\sum_{x,y \in \Omega} x \cdot T(x, y)}{\sum_{x,y \in \Omega} T(x, y)}, \quad y_c = \frac{\sum_{x,y \in \Omega} y \cdot T(x, y)}{\sum_{x,y \in \Omega} T(x, y)}, \quad (3.4)$$

и переместить центр тяжести в начало координат:

$$x^* = x - x_c, \quad y^* = y - y_c, \quad (3.5)$$

где Ω - двумерная область определения изображения объекта, $T(x, y)$ – яркость пиксела с координатами $\{x, y\}$. Т.е. все пикселы, принадлежащие изображению объекта, перемещаются на $-x_c$ и на $-y_c$, соответственно.

В ходе дальнейших вычислений вместо реального перемещения пикселов можно учесть, что координаты x и y каждого пиксела, подставляемые в

ниже используемые формулы, должны быть представлены как $x - x_c$ и $y - y_c$.

1.7. В соответствии с [1, 2] направление Θ и величину μ сжатия изображения, центрированного согласно пункту 1.6, можно вычислить с использованием следующих выражений:

$$\mu = \pm \sqrt{\frac{D + \sqrt{C^2 + B^2}}{D - \sqrt{C^2 + B^2}}}, \quad (3.6)$$

$$\Theta = 0,5 \arctg(C/B) + \pi n, \quad (3.7)$$

$$B = \iint_{\Omega} T(r, \varphi) r^2 \cos(2\varphi) dr d\varphi, \quad (3.8)$$

$$C = \iint_{\Omega} T(r, \varphi) r^2 \sin(2\varphi) dr d\varphi, \quad (3.9)$$

$$D = \iint_{\Omega} T(r, \varphi) r^2 dr d\varphi, \quad (3.10)$$

где $T(r, \varphi)$ – яркость пиксела изображения объекта, представленного в полярной системе координат, Ω – область определения изображения объекта в полярной системе координат $\{r, \varphi\}$ – соответственно радиус и азимут пиксела в полярной системе координат.

Используя тригонометрические формулы $\cos(2\varphi) = \cos^2(\varphi) - \sin^2(\varphi)$ и $\sin(2\varphi) = 2\cos(\varphi)\sin(\varphi)$ и переходя от интегралов к дискретным суммам, а также принимая во внимание, что теперь мы имеем дело с изображением объекта, центрированным согласно формулам (3.5), выражения (3.8), (3.9) и (3.10) можно представить как

$$B = \sum_{x, y \in \Omega} T(x, y) [(x - x_c)^2 - (y - y_c)^2], \quad (3.11)$$

$$C = \sum_{x, y \in \Omega} T(x, y) 2(x - x_c)(y - y_c), \quad (3.12)$$

$$D = \sum_{x, y \in \Omega} T(x, y) [(x - x_c)^2 + (y - y_c)^2]. \quad (3.13)$$

Подставив в формулы (3.6) и (3.7) значения B , C и D , вычисленные согласно формулам (3.11), (3.12) и (3.13), можно рассчитать направление Θ и величину μ сжатия центрированного изображения и, применив к нему масштабирование в $1/\mu$ раз в направлении Θ , компенсировать преобразование \mathbf{A}_3 , упомянутое выше в формуле (3.3). Такое

преобразование координат пикселей можно описать следующим выражением:

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \mathbf{R}(+\Theta) \mathbf{S}(1/\mu) \mathbf{R}(-\Theta) \begin{pmatrix} x^* \\ y^* \end{pmatrix},$$

где x^+ и y^+ – абсцисса и ордината пикселей центрированного изображения после компенсации его масштабирования с коэффициентом μ в направлении Θ , x^* и y^* – координаты пикселей центрированного изображения, $\mathbf{R}(-\Theta)$ – матрица вращения изображения на угол $-\Theta$ (поворачивает направление масштабирования к оси абсцисс), $\mathbf{S}(1/\mu)$ – матрица масштабирования в $1/\mu$ раз в направлении оси абсцисс, $\mathbf{R}(+\Theta)$ – матрица обратного вращения изображения на угол $+\Theta$.

$$\mathbf{R}(+\Theta) = \begin{pmatrix} \cos \Theta & -\sin \Theta \\ \sin \Theta & \cos \Theta \end{pmatrix},$$

$$\mathbf{R}(-\Theta) = \begin{pmatrix} \cos(-\Theta) & -\sin(-\Theta) \\ \sin(-\Theta) & \cos(-\Theta) \end{pmatrix},$$

$$\mathbf{S}(1/\mu) = \begin{pmatrix} 1/\mu & 0 \\ 0 & 1 \end{pmatrix}.$$

Тогда

$$x^+ = \frac{1}{\mu} [(x - x_c) \cos(-\Theta) - (y - y_c) \sin(-\Theta)] \cos \Theta - [(x - x_c) \sin(-\Theta) + (y - y_c) \cos(-\Theta)] \sin \Theta, \quad (3.14)$$

$$y^+ = \frac{1}{\mu} [(x - x_c) \cos(-\Theta) - (y - y_c) \sin(-\Theta)] \sin \Theta + [(x - x_c) \sin(-\Theta) + (y - y_c) \cos(-\Theta)] \cos \Theta. \quad (3.15)$$

1.8. Теперь остаются не измеренными и не компенсированными преобразования \mathbf{A}_5 , \mathbf{A}_6 и \mathbf{A}_4 равномерного масштабирования, вращения и зеркального отражения (см. выше в п. 1.5). Преобразование \mathbf{A}_5 не зависит от изменения углового положения пикселя и не влияет на \mathbf{A}_6 и \mathbf{A}_4 , поэтому параметры $\mathbf{A}_5 = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix}$ можно измерить по проекции изображения на радиус полярной системы координат. В соответствии с [1, 2] коэффициент M равномерного масштабирования изображения может быть измерен согласно формуле (3.16) по смещению центра тяжести его проекции на радиус полярной системы координат относительно эталонного положения K центра тяжести этой проекции,

если изображение было предварительно центрировано (см. выше в п. 1.6), и если его пространственно неравномерное масштабирование было компенсировано согласно п. 1.7.

$$M = \left\{ \frac{\sum_{r, \varphi \in \Omega} r \cdot T(r, \varphi)}{K \cdot \sum_{r, \varphi \in \Omega} T(r, \varphi)} \right\} = \left\{ \frac{\sum_{x, y \in \Omega} T(x, y) \cdot \sqrt{(x^+)^2 + (y^+)^2}}{K \cdot \sum_{x, y \in \Omega} T(x, y)} \right\}. \quad (3.16)$$

Примем для определенности $K=10$, тогда, подставив в (3.16) координаты x^+ и y^+ пикселей исходного изображения, полученные в результате компенсации преобразований \mathbf{A}_1 , \mathbf{A}_2 и \mathbf{A}_3 по формулам (3.4), (3.5), (3.14) и (3.15), вычислим коэффициент M равномерного масштабирования изображения, что позволит затем компенсировать преобразование \mathbf{A}_5 , выполнив операцию

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} 1/M & 0 \\ 0 & 1/M \end{pmatrix} \begin{pmatrix} x^+ \\ y^+ \end{pmatrix}. \quad (3.17).$$

1.9. Остались не компенсированными преобразования \mathbf{A}_6 и \mathbf{A}_4 вращения и зеркального отражения. Эти преобразования влияют на равномерное изменение угловой координаты пикселей в полярной системе координат и не зависят от распределения пикселей по радиальной координате. Поэтому параметры этих преобразований могут быть измерены по проекции изображения на окружность. Однако общий подход к измерению параметров этих преобразований требует применения спектрального анализа [1, 2]. При этом не удастся оперировать только численными значениями координат пикселей, как это делалось в предыдущих пунктах. Вместо этого придется реально переводить изображение в полярную систему координат, проецировать пиксели изображения на окружность и использовать дискретное преобразование Фурье. Эти операции исследуются в лабораторных работах №2, №4 и №5, поэтому в настоящей работе измерение параметров преобразований \mathbf{A}_6 и \mathbf{A}_4 вращения и зеркального отражения производиться не будет.

В заключение необходимо отметить, что описанный выше в пунктах 1.6-1.8 математический аппарат оперирует с изображениями объектов, помещенными на нулевой фон, поэтому описанная методика измерения параметров и компенсации аффинного преобразования пригодна только в ситуациях, когда имеются средства опознавания на изображении пикселей, принадлежащих анализируемому объекту.

2 Порядок выполнения лабораторной работы

- 2.1.** Получить у преподавателя изображения для выполнения лабораторной работы. Необходимо помнить, что для корректности выполнения выше описанных процедур измерения параметров аффинного преобразования изображения объекта, оно должно быть отделено от фона и других объектов; именно такие изображения объектов на нулевом фоне предлагаются студентам для экспериментов.
- 2.2.** Измерить параметры аффинных преобразований этих изображений и нормализовать их относительно этих преобразований по методике, изложенной в разделе 1. Нормализующие геометрические преобразования с вычисленными параметрами можно выполнить, используя программу, написанную для лабораторной работы №2, или вращая и масштабируя изображения отдельно по осям x и y с помощью подходящих графических редакторов. Компенсацию выше рассмотренных компонентов аффинного преобразования можно также выполнить, пересчитывая координаты каждого пиксела по формулам (3.14), (3.15) и (3.17). Это будет не вполне корректно, т.к. при этом не будет выполняться интерполяция, в то время как пересчитанные координаты пикселей не будут целыми числами, и их придется округлять до целого, однако, для предложенного эталонного значения коэффициента равномерного масштабирования, соответствующего $K=10$, эта некорректность не скажется на внешнем виде нормализованных изображений.
- 2.3.** Проанализировать и объяснить результаты нормализации изображений.
- 2.4.** Представить полученные результаты преподавателю.

3 Вопросы для самопроверки

1. Какие физические эффекты являются источником появления проективных преобразований в изображениях, анализируемых в системах компьютерного зрения?
2. В каких случаях проективное преобразование можно аппроксимировать аффинным?
3. В виде какой последовательности простых геометрических трансформаций можно представить аффинное преобразование изображения?

Литература

1. Вадим Луцив. Автоматический анализ изображений. Объектно-независимый структурный подход. – Саарбрюккен, Германия: Ламберт Академик Пабблишинг, 2011. – 300 с.
2. Луцив В.Р. Метод обработки изображений при управлении роботами в ГАП. – В кн.: ЭВМ в проектировании и производстве: Сб. статей. Вып.2 /Под общ. ред. Г.В. Орловского. – Л.: Машиностроение, Ленингр. отделение, 1985. – с.154-162.

3. Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А. С. Потапов, М. В. Петерсон, В. А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.
4. Bradski G. and Kaehler A. Learning OpenCV. Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media, Inc., 2008. 556 p.

ЛАБОРАТОРНАЯ РАБОТА №4

Распознавание объектов на изображении методом Фурье-Меллина

1 Теоретическая часть

Одним из простейших и наиболее распространенных путей поиска эталонного образа в исследуемом изображении является вычисление функции их взаимной корреляции $C(x,y)$:

$$C(x,y) = (f \otimes g)_{x,y} = \sum_{k,l} f_{x+k,y+l}^* g_{k,l},$$

где $f_{x,y}$ и $g_{k,l}$ – значения яркости пикселей сопоставляемых исследуемого и эталонного изображений, соответственно, x и y – декартовы координаты пикселя в исследуемом изображении или в двумерном поле вычисленной функции взаимной корреляции (номер строки и номер столбца пикселей, соответственно), $*$ – операция комплексного сопряжения (ее можно опустить, т.к. в данном случае пикселям приписаны действительные значения яркости), k и l – декартовы координаты пикселя в эталонном изображении. Путем поиска максимума построенной таким образом двумерной корреляционной функции можно определить местонахождение искомого эталона (объекта) в исследуемом изображении. Значение корреляционной функции в точке ее максимума характеризует степень сходства сопоставляемых изображений.

Сопоставление изображений путем вычисления функции их взаимной корреляции как частный случай применения согласованного фильтра является оптимальным решением при высоких уровнях аддитивной шумовой составляющей, отличающей изображение искомого объекта от эталонного. Однако корреляционный отклик разрушается, если взаимное геометрическое преобразование сопоставляемых изображений описывается более сложными трансформациями системы координат, чем простые взаимные линейные смещения вдоль декартовых координатных осей. В то же время для некоторых видов преобразования системы координат изображения эта проблема может быть решена. В этой лабораторной работе рассматривается, каким образом можно применять корреляционную функцию для сопоставления изображений, имеющих взаимные масштабные преобразования, вращения и линейные смещения вдоль декартовых осей координат.

Для достижения поставленной цели изображение преобразуется из декартовой в полярно-логарифмическую систему координат, как это было объяснено выше в описании лабораторной работы №2. В этом случае координаты пикселя на плоскости описываются длиной R радиуса-вектора \mathbf{R} , соединяющего центр пикселя с началом координат, и направлением радиуса-вектора, задаваемым углом ϕ азимута, которые вычисляются по формулам (4.1) и (4.2), соответственно.

$$\phi = \arctg(y/x); \quad (4.1)$$

$$R = \ln[(x^2 + y^2)^{1/2}]. \quad (4.2)$$

Масштабирование исходного изображения, заданного в декартовой системе координат, приводит в новой системе координат к линейному смещению его образа вдоль радиальной оси R , а вращение исходного изображения вокруг начала декартовой системы координат приводит к линейному циклическому смещению его образа, представленного в полярно-логарифмической системе координат, вдоль оси азимута ϕ . Следовательно, если два изображения, которые мы хотим сопоставить, взаимно масштабированы и повернуты относительно начала декартовой системы координат, но не претерпели в этой системе координат линейных смещений, они могут быть успешно сопоставлены с помощью вычисления их функции взаимной корреляции в полярно-логарифмической системе координат. При этом антилогарифм радиальной координаты пика корреляции покажет величину взаимного масштабирования, а азимутальная координата пика будет соответствовать величине угла взаимного вращения сопоставляемых изображений.

Имеется, однако, возможность применить корреляционный метод сопоставления и для изображений, претерпевших взаимное вращение, масштабирование и линейные сдвиги на плоскости. Для этого оба сопоставляемых изображения сначала подвергаются преобразованию Фурье, а переводу в полярно-логарифмическую систему координат подвергается амплитудная часть $A(\omega_x, \omega_y)$ их пространственного Фурье-спектра, показанная в формуле (4.3) на примере входного изображения $f(x, y)$ размерами $N \times M$ пикселей.

$$A(\omega_x, \omega_y) = \text{abs} \left[\sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \exp(-j2\pi x \omega_x / N) \exp(-j2\pi y \omega_y / M) \right], \quad (4.3)$$

где j – мнимая единица, x, y – номера столбцов и строк пикселей, а ω_x и ω_y – номера отсчетов пространственной частоты вдоль соответствующих координатных осей. Согласно свойствам преобразования Фурье модули Фурье-гармоник пространственного спектра изображения, вычисленного в декартовой системе координат, инвариантны к сдвигу изображения вдоль декартовых координатных осей, поэтому модули Фурье-образов двух идентичных изображений, преобразованные в полярно-логарифмическую систему координат, будут отличаться только взаимными смещениями вдоль осей азимута и радиуса. Следовательно, для определения параметров взаимного вращения и масштабирования сопоставляемых изображений достаточно будет измерить координаты пика функции взаимной корреляции модулей их Фурье-образов, преобразованных в полярно-логарифмическую систему координат. Высота пика рассчитанной таким образом функции взаимной корреляции будет нести информацию о сходстве или несходстве

сопоставляемых изображений. Для не соответствующих друг другу сопоставляемых изображений функция взаимной корреляции не будет иметь явно выраженных пиков, а если соответствующие друг другу сопоставляемые изображения подверглись более сложному взаимному преобразованию, чем масштабирование, вращение и сдвиги (упрощенная группа подобия, не включающая преобразование зеркальным отражением), пик корреляционной функции будет размываться все сильнее по мере отклонения преобразования от упрощенной группы подобия и в конечном итоге «утонет в шуме». Аналогичная деградация пика функции корреляции будет происходить при недостаточном взаимном перекрытии сопоставляемых изображений, вызванном, например, значительными их взаимными сдвигами (когда большая часть одного изображения не содержится в другом), или существенным взаимным масштабированием.

Для автоматического сопоставления анализируемого изображения с эталонными не обязательно выполнять поиск пика функций их взаимной корреляции, как это делалось выше в лабораторных работах №1 и №2. Достаточно, например, сравнить величины коэффициентов корреляции анализируемого изображения с разными эталонными и считать, что анализируемое изображение соответствует тому эталону, для которого модуль коэффициента корреляции максимален и превышает некоторый пороговый уровень. При этом коэффициент C_m корреляции изображения с m -тым эталоном может быть вычислен по формуле (4.4).

$$C_m = \frac{\sum_{j=1}^N (B_j - \bar{B})(B_j^{(m)} - \bar{B}^{(m)})}{\sqrt{\left(\sum_{j=1}^N (B_j - \bar{B})^2 \right) \left(\sum_{j=1}^N (B_j^{(m)} - \bar{B}^{(m)})^2 \right)}}, \quad (4.4)$$

где B_j – яркость j -того пиксела анализируемого образа, $B_j^{(m)}$ – яркость j -того пиксела m -того эталонного образа, \bar{B} и $\bar{B}^{(m)}$ – среднее значение (математическое ожидание) яркости пикселов анализируемого образа и m -того эталонного образа, соответственно, N – число пикселов. Однако для сопоставления изображений с помощью вычисленного таким образом коэффициента корреляции число N пикселов анализируемого и число $N^{(m)}$ пикселов эталонного образа должны совпадать. И более того, такое сопоставление возможно, только если сопоставляемые образы инвариантны к преобразованиям, характерным для решаемой задачи, – в нашем случае это – преобразования смещениями, вращением и масштабированием. Как это было описано выше, образ, состоящий из амплитуд Фурье-гармоник, представленный в полярно-логарифмической системе координат, инвариантен к смещениям сопоставляемого изображения вдоль декартовых координатных осей, но этот образ сдвигается в полярно-логарифмической

системе координат вдоль осей абсцисс и ординат под действием масштабирования и вращения исходного изображения. Чтобы анализировать образ, инвариантный к таким сдвигам, нужно повторно вычислить от этого образа двумерное преобразование Фурье и взять модули Фурье гармоник. Такой образ будет инвариантен к смещениям, вращению и масштабированию исходного изображения.

Для того чтобы уравнивать числа сопоставляемых пикселей анализируемого и эталонного изображения (сделать $N = N^{(m)}$ в формуле (4.4)), необходимо в полярно-логарифмической системе координат перед вычислением преобразования Фурье сопоставляемых образов увеличить число пикселей меньшего из них по радиальной координате до числа пикселей большего из них (дополнить снизу меньший образ пикселями, соответствующим средним значениям яркости этого образа в полярно-логарифмической системе координат).

При вычислении быстрого преобразования Фурье надо не забыть дополнить размеры сопоставляемых образов по каждой координате до значений, соответствующих ближайшей большей или равной увеличиваемому размеру величине двойки в целой степени, – в соответствии с используемым алгоритмом (рекомендуется алгоритм Кули-Тьюки). При использовании других алгоритмов быстрого преобразования Фурье (например, алгоритма Винограда) надо не забыть увеличить размер изображения до величины, требуемой используемым алгоритмом. Если для выполнения преобразования Фурье двумерного образа используется одномерный алгоритм, его можно сперва применить последовательно к каждой строке пикселей образа. Затем для преобразованному таким путем двумерного образа одномерное преобразование Фурье применяется последовательно к каждому столбцу. Если используемый алгоритм вычисляет двумерное преобразование Фурье, его можно одновременно применять ко всему двумерному массиву пикселей преобразуемого образа. Готовые программные коды, реализующие преобразование Фурье можно найти в интернете. Алгоритмы быстрого преобразования Фурье приводятся также в книгах по цифровой обработке сигналов (см., например, Л. Рабинер и Б. Гоулд. Теория и применение цифровой обработки сигналов / Пер. с англ. под ред. Ю.Н. Александрова. – М.: Мир, 1978. – 848 с.).

подавляющее большинство доступных программ быстрого преобразования Фурье принимает на входе массивы комплексных чисел – действительную и мнимую части комплексного числа для каждого пиксела – и выдает результат также в виде массива комплексных чисел, состоящих из действительной и мнимой частей. При вычислении модуля представленного таким образом комплексного числа берется евклидова сумма его действительной и мнимой частей.

При выполнении данной лабораторной работы преобразование двумерного образа из декартовой системы координат в полярно-логарифмическую можно выполнить так, как это было описано в лабораторной работе №2. Можно

найти и готовые программы такого преобразования, написанные в среде MATLAB или на языке С в одной из версий библиотеки OpenCV. Однако, пользуясь готовыми программными моделями, рекомендуется убедиться в корректности их работы. В отличие от алгоритма преобразования, предложенного в лабораторной работе №2, некоторые из готовых программ выполняют преобразование в полярно-логарифмическую систему координат без интерполяции. Это вносит ошибки в процесс преобразования и в конечном итоге увеличивает мощность шума в вычисляемом корреляционном поле и размывает корреляционный максимум, что делает плохо различимыми корреляционные коэффициенты, вычисленные для анализируемого изображения с соответствующими и не соответствующими ему эталонными образами.

2 Порядок выполнения лабораторной работы

1. Для выполнения работы студентам предлагается самостоятельно выбрать три разных изображения размером не меньше 512 X 512 пикселей и взять эти изображения в качестве эталонных образов. Взять одно из этих изображений также в качестве исследуемого (распознаваемого) – вырезать это изображение и соответствующий ему эталон из одного и того же места некоторого исходного изображения больших размеров.
2. Вычислить коэффициенты корреляции исследуемого изображения с каждым из эталонных, как это было описано выше:
 - а) Вычислить двумерное преобразование Фурье каждого изображения и в каждый пиксел преобразованного образа записать модуль преобразования Фурье. Пример Фурье спектра (амплитуды гармоник) изображения приведен на рис. 4.1. Необходимо обратить внимание, что программы быстрого преобразования Фурье (согласно правилам вычисления периодического спектра дискретизированных сигналов) часто располагают низкочастотные гармоники не в центре изображения, а в его углах, нулевая гармоника – в левом верхнем углу.
 - б) Перевести образы, вычисленные согласно пункту (а) в полярно-логарифмическую систему координат (в отличие от лабораторной работы №2 рекомендуется для облегчения вычисления быстрого преобразования Фурье взять для каждого образа в полярно-логарифмической системе координат 256 или 512 отсчетов по угловой координате).
 - в) Дополнить образы всех участвующих в экспериментах изображений, вычисленные согласно пункту (б), до одинаковых размеров, подходящих для вычисления быстрого преобразования Фурье, по радиальной координате.
 - г) Вычислить двумерное преобразование Фурье от каждого образа, полученного согласно пункту (в), и в каждый пиксел преобразованного образа записать модуль преобразования Фурье.

- д) Вычислить коэффициент корреляции исследуемого образа с каждым из эталонных. Оценить, как изменяется коэффициент корреляции, если эталонный образ соответствует исследуемому образу («правильный эталон») или не соответствует ему.

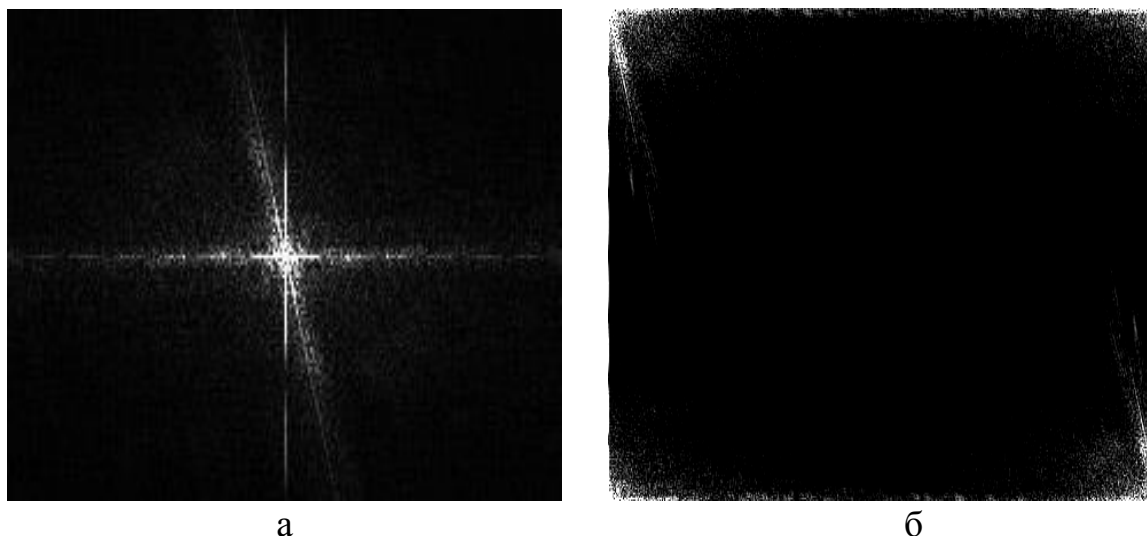


Рисунок 4.1 – Амплитудная часть пространственного Фурье-спектра изображения: а – традиционная форма представления спектра (нулевая гармоника в центре его изображения), б – форма, вычисляемая программами дискретного преобразования Фурье (нулевая гармоника в левом верхнем углу)

3. Взять те же соответствующие друг другу исследуемое и эталонное изображения, которые рассматривались в пункте 2. Изменить масштаб исследуемого изображения относительно эталонного приблизительно на 20% и, выполняя для рассматриваемой пары изображений операции, перечисленные в пункте 2, определить, как в результате выполненного масштабирования изменился коэффициент корреляции исследуемого образа с «правильным эталоном».
4. Выполнить операции, соответствующие пункту 3, для взаимного размасштабирования сопоставляемых изображений на 50% и на 100%. Оценить, как изменяется при этом значение коэффициента корреляции, можно ли будет при значительном масштабировании соотнести исследуемое изображение с «правильным эталоном» по значению коэффициента корреляции.
5. Выполнить операции пункта 3, исследуя, как влияет на величину коэффициента корреляции степень взаимного перекрытия исследуемого изображения и «правильного эталона» при их взаимном сдвиге вдоль декартовых осей координат. Для этого при подготовке сопоставляемых изображений согласно пункту 1 исследуемое изображение вырезать из исходного большего по размерам изображения со смещением по оси абсцисс на 20% (относительно горизонтального размера исследуемого изображения), 40% и 60% по сравнению с положением «правильного» эталонного изображения.

3 Вопросы для самопроверки

1. С помощью какого преобразования изображения получается образ, инвариантный к сдвигам?
2. В каких случаях результаты сопоставления изображений методом Фурье-Меллина перестают быть устойчивы по отношению к их взаимному масштабированию и сдвигам?
3. По какому показателю сходства можно сравнивать описания изображений, инвариантные к вращению, масштабированию и сдвигам?

Литература

1. Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А. С. Потапов, М. В. Петерсон, В. А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.
2. Bradski G. and Kaehler A. Learning OpenCV. Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media, Inc., 2008. 556 p.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. 1072 с.

ЛАБОРАТОРНАЯ РАБОТА №5

Распознавание изображений, подвергнутых аффинным преобразованиям

1 Теоретическая часть

Для описания зрительных систем высших животных и человека и большинства систем компьютерного зрения могут быть применены схемы, представленные на рис.3.1.а, рис.3.1.б и рис.3.1.в и в деталях рассмотренные выше в материалах лабораторной работы №3 (для успешного выполнения настоящей лабораторной работы рекомендуется с материалами работы №3 повторно подробно ознакомиться). В указанных материалах рассматривался частный случай, когда точки плоской поверхности объекта (что для большинства объектов реального мира выполняется по крайней мере локально) проецируются на плоскую матрицу светочувствительных элементов (что, как правило, соответствует истине, если мы имеем дело с системами компьютерного зрения, и справедливо для локальных участков сетчатки глаза). При этом, если плоскость, содержащая светочувствительные элементы, не параллельна отображаемой поверхности объекта, но его размеры намного меньше расстояния до объектива камеры, геометрическое преобразование (искажение) изображения объекта, сформированного на световоспринимающей матрице камеры, относится к группе аффинных, которые описываются выражением

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \Leftrightarrow \begin{cases} x' = a_1x + a_2y + a_3 \\ y' = a_4x + a_5y + a_6 \end{cases}, \quad (5.1)$$

где x, y – абсцисса и ордината точки исходного изображения, x', y' – преобразованные координаты этой точки, $a_1 \dots a_6$ – параметры преобразования.

Как было показано в материалах лабораторной работы №3, геометрическое преобразование (5.1) можно представить в векторно-матричной форме в виде последовательности (5.2) более простых преобразований.

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4 \mathbf{A}_5 \mathbf{A}_6, \quad (5.2)$$

где \mathbf{A}_1 соответствует линейному смещению центра тяжести изображения вдоль оси абсцисс, \mathbf{A}_2 соответствует смещению центра тяжести изображения вдоль оси ординат, \mathbf{A}_3 соответствует сжатию вдоль некоторого направления, \mathbf{A}_4 соответствует зеркальному отражению, \mathbf{A}_5 соответствует масштабированию, равномерному по двум декартовым координатам, \mathbf{A}_6 соответствует вращению объекта в плоскости изображения относительно начала координат.

В материалах лабораторной работы №3 было также показано, что вычислить вектор $(x'', y'', 1)^T$ исходных (не искаженных) координат пикселей (компенсировать их аффинное преобразование) можно с помощью следующей последовательности операций:

$$\begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} = \mathbf{A}_6^{-1} \mathbf{A}_5^{-1} \mathbf{A}_4^{-1} \mathbf{A}_3^{-1} \mathbf{A}_2^{-1} \mathbf{A}_1^{-1} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \equiv \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \quad (5.3)$$

В работе №3 приведены также математические выражения, с помощью которых можно выполнить измерение параметров и компенсацию (5.3) аффинного преобразования исследуемых изображений. Применять эти выражения для измерения параметров аффинного преобразования можно только, подставляя в них яркость и координаты пикселей относящихся непосредственно к изображению конкретного выбранного объекта, поэтому перед компенсацией аффинного преобразования изображение такого объекта должно быть выделено из фона (помещено на нулевой фон).

В настоящей лабораторной работе исследуются методы автоматического распознавания изображений, подвергнутых аффинным преобразованиям (описание неизвестного распознаваемого изображения сопоставляется с описаниями эталонных изображений, принадлежащих разным известным классам). Процедуре распознавания предшествуют рассмотренные выше процедуры измерения параметров и компенсация аффинного преобразования анализируемого изображения (нормализации изображения) для приведения его в эталонное состояние. После компенсации преобразования изображение описывается вектором признаков. Аналогичной нормализации и описанию векторами признаков должны быть подвергнуты эталонные изображения – представители известных классов. Распознавание анализируемого неизвестного изображения (классификация выделенного в нем объекта) выполняется путем сопоставления описывающего его вектора признаков с векторами признаков эталонных изображений. Действия, которые необходимо выполнить для нормализации, описания и сопоставления распознаваемого и эталонных изображений приводятся в пунктах, приведенных ниже.

1.1. В этой лабораторной работе предполагается, что аффинное преобразование, которому подвергались анализируемые изображения, описываются следующей цепочкой простых операций:

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_5 \mathbf{A}_6. \quad (5.4)$$

Операция \mathbf{A}_4 зеркального отражения, входившая в состав полного аффинного преобразования согласно формуле (5.2), из рассмотрения в

(5.4) исключена, т.к. на практике она встречается редко. Измерение параметров и компенсацию простых преобразований A_1 , A_2 , A_3 и A_5 предлагается выполнить согласно методике и формулам, приведенным в лабораторной работе №3. Для упрощения вычислений студентам разрешается выполнять операции компенсации аффинного преобразования путем простого пересчета координат пикселей согласно формулам (3.14), (3.15) и (3.17), как это было описано в материалах лабораторной работы №3. Использование в формуле (3.17) эталонного значения коэффициента масштабирования $K=10$ (см. в материалах лабораторной работы №3) позволяет выполнять геометрические преобразования изображений без интерполяции, но обеспечивает такие размеры нормализованного изображения, которые недостаточны для выполнения операций, описанных ниже в пунктах 1.2 и 1.3. В связи с этим перед выполнением дальнейших операций студентам предлагается увеличить в пять раз размеры изображений, нормализованных относительно аффинного преобразования выше описанным образом. Такое увеличение можно выполнить любым подходящим графическим редактором (например, Adobe Photoshop или GIMP), или сделать это в собственной программе путем пятикратного повторения каждого столбца и каждой строки нормализованного изображения.

1.2. Теперь остается некомпенсированным преобразование A_6 вращения изображения относительно начала координат. Вместо нормализации изображений относительно этого преобразования предлагается описывать распознаваемые и эталонные изображения векторами признаков, инвариантными к вращению. Такие описания строятся для каждого распознаваемого и эталонного изображения, нормализованных представленным выше способом, путем выполнения следующей последовательности операций:

1.2.1. Нормализованные изображения, для которых яркость пикселей задавалась в декартовой системе координат $\{x, y\}$ как $f(x, y)$, переводятся в полярную систему координат $\{R, \varphi\}$: $f(x, y) \rightarrow f(R, \varphi)$. Предполагается при этом, что центр полярной системы координат совпадает с центральным пикселем исходно заданного для анализа изображения, в который также перенесен в результате выше описанной нормализации центр тяжести анализируемого на изображении объекта. Азимут φ пикселя вычисляется как арктангенс отношения ординаты y пикселя к его абсциссе x .

1.2.2. Диапазон $[0; 2\pi)$ значений азимутальной координаты φ пикселей разделяется на 64 следующих друг за другом поддиапазонов шириной $2\pi/64$ каждый. Каждому из 64 азимутальных поддиапазонов на изображении ставится в соответствие сектор, имеющий вершину в центре изображения и ограниченный исходящими из центра лучами, направленными вдоль границ поддиапазона. Будем считать, что

пиксел принадлежит некоторому выбранному сектору (азимутальному поддиапазону) с номером n ($n=0\dots63$), если его азимутальная координата φ больше или равна азимутальной координате $\varphi_n=2\pi n/64$ нижней границы этого поддиапазона и меньше координаты $\varphi_{n+1}=2\pi(n+1)/64$ его верхней границы. Внутри каждого сектора подсчитывается сумма значений попадающих в него пикселов, полученная для каждого сектора сумма делится на сумму значений пикселов всего изображения. Нормированные выше описанным способом суммы значений пикселов внутри отдельных секторов сохраняются как элементы массива признаков $f'(n)$, описывающего изображение объекта, как это показано на рис. 5.1.

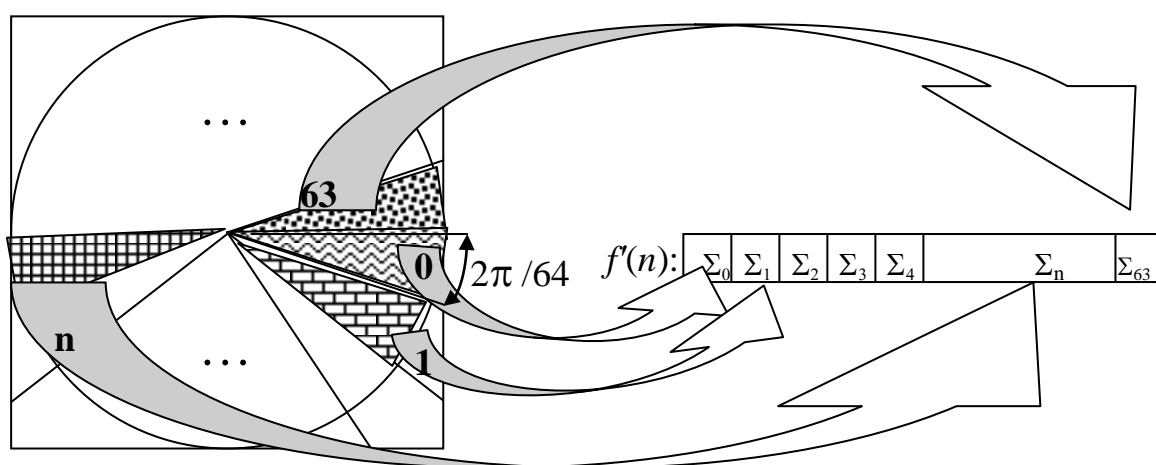


Рисунок 5.1 – Формирование массива признаков как нормированных сумм значений пикселов внутри 64 секторов, имеющих общий центр в центральном пикселе изображения.

1.2.3. Вращение объекта в плоскости изображения приводит к циклическому сдвигу элементов массива признаков $f'(n)$, полученного согласно пунктам 1.2.1 и 1.2.2, поэтому этот массив должен теперь быть преобразован в форму, инвариантную к сдвигу. Этого можно достичь, если использовать в качестве признаков изображения амплитуды $A(\omega)$ гармоник спектра (5.5), рассчитанные путем выполнения дискретного преобразования Фурье над элементами массива $f'(n)$, сформированного согласно пунктам 1.2.1 и 1.2.2. Запишем вычисленный таким образом пространственный спектр в виде вектора признаков \mathbf{A} , действительными компонентами которого будут модули дискретных спектральных отсчетов:

$$\mathbf{A} \leftrightarrow A(\omega) = \text{abs} \left[\sum_{n=0}^{63} f'(n) \exp(-j2\pi n \omega / 64) \right], \quad (5.5)$$

где j – мнимая единица, $\omega=0\dots 63$ – номера дискретных отсчетов пространственной частоты вдоль азимутальной координаты φ . Преобразование Фурье рекомендуется выполнять с использованием алгоритма Кули-Тьюки – самого распространенного алгоритма быстрого преобразование Фурье (БПФ). Как раз для использования этого алгоритма и был сформирован массив признаков длиной $64=2^6$. Каждый из 64 действительных компонентов вектора спектральных признаков получается путем вычисления евклидовой суммы (корень квадратный из суммы квадратов слагаемых) соответствующих друг другу (имеющих один и тот же номер спектральной гармоники) действительной и мнимой компонент спектра, рассчитанного с помощью БПФ.

Таким образом, в результате выполнения операций, описанных в пунктах 1.2.1 – 1.2.3 настоящего раздела, строятся вектор спектральных признаков \mathbf{A}_0 распознаваемого изображения и N векторов спектральных признаков \mathbf{A}_i ($i=1\dots N$) эталонных изображений, инвариантные к геометрическому преобразованию (5.4).

1.3. Теперь для сопоставления 64-компонентного вектора признаков \mathbf{A}_0 изображения распознаваемого объекта, вычисленного согласно пунктам 1.1 и 1.2, с N эталонными векторами признаков \mathbf{A}_i ($i=1\dots N$), вычисленными аналогичным способом, выполняются следующие операции:

1.3.1. Для каждого (i -того) эталонного вектора признаков \mathbf{A}_i ($i=1\dots N$) вычисляется согласно евклидовой метрике модуль разности с вектором признаков \mathbf{A}_0 распознаваемого изображения:

$$\Delta_i = \|\mathbf{A}_0 - \mathbf{A}_i\| = \sqrt{\sum_{\omega=0}^{63} [A_0(\omega) - A_i(\omega)]^2}.$$

1.3.2. Рассматриваемое изображение неизвестного объекта относится к тому (J -тому) классу известных объектов, для которого разность Δ_j имеет минимальное значение среди всех Δ_i ($i=1\dots N$), вычисленных согласно пункту 1.3.1:

$$J = \arg \min_{i=1\dots N} (\Delta_i).$$

2 Порядок выполнения лабораторной работы

2.1. Для выполнения лабораторной работы потребуются три «эталонных» изображения, принадлежащие к трем разным классам, и одно «неизвестное подлежащее распознаванию» изображение, относящееся к одному из трех указанных классов. Необходимо еще раз напомнить, что

нормализация изображений относительно аффинного преобразования, их инвариантное описание и распознавание путем сопоставления с эталонными векторами признаков, как это описано в разделе 1, могут выполняться корректно, только если изображения анализируемого и эталонных объектов предварительно будут отделены от фона (пиксели, относящиеся к представляющему интерес объекту, остаются неизменными, пиксели, к нему не относящиеся, обнуляются). Приготовить изображения, необходимые для экспериментов, можно двумя путями:

- 2.1.1. Использовать в качестве «эталонных» три изображения, полученные от преподавателя при выполнении лабораторной работы №3. Это – изображения простых геометрических фигур, уже помещенных на нулевой (черный) фон. Для формирования «неизвестного» подлежащего автоматической классификации изображения взять одно из приготовленных выше описанным способом эталонных и повернуть его в любом подходящем графическом редакторе (например, Adobe Photoshop или GIMP) на 20° – 30° . Проследить при этом, чтобы в результате выполнения вращения в изображении не появились дополнительные области ненулевой яркости. Для единообразного использования в разрабатываемой программной модели функций чтения графических файлов рекомендуется сохранить повернутое таким образом изображение в файле того же стандартного графического формата, что и файлы приготовленных «эталонных» изображений.
- 2.1.2. При желании вместо изображений, полученных согласно пункту 2.1.1, студенты могут приготовить самостоятельно три «эталонных» и одно «неизвестное» изображение, получаемое путем вращения одного из «эталонных», и сохранить эти изображения в файлах одного из стандартных графических форматов. При этом необходимо помнить, что приготовленные таким образом изображения объектов должны быть помещены на нулевой фон, а размеры этих объектов должны приблизительно соответствовать размерам объектов в изображениях, полученных согласно пункту 2.1.1. В противном случае геометрическое преобразование изображения, выполняемое согласно пункту 1.1 без интерполяции, может стать некорректным, а вместо пятикратного увеличения нормализованного изображения, рекомендуемого в том же пункте, может потребоваться выбирать подходящий коэффициент увеличения самостоятельно.
- 2.2. Нормализовать четыре изображения, приготовленные согласно пункту 2.1, относительно геометрических преобразований, используя методику, описанную выше в пункте 1.1, и сохранить нормализованные изображения в файлах одного из стандартных графических форматов, чтобы обеспечить возможность их визуальной интерпретации.

- 2.3. Вычислить для четырех нормализованных изображений (одного «неизвестного» и трех «эталонных представителей известных классов образов») инвариантные векторы признаков согласно методике, изложенной в пункте 1.2.
- 2.4. Выполнить классификацию «неизвестного» изображения путем сопоставления его инвариантного вектора признаков с «эталонными векторами признаков» трех известных классов согласно методике, изложенной в пункте 1.3.
- 2.5. Проанализировать и объяснить полученные в пунктах 2.2–2.4 результаты.
- 2.6. Представить полученные результаты преподавателю.

3 Вопросы для самопроверки

1. Относительно каких простых геометрических преобразований, составляющих аффинную группу, нормализуются эталонные и распознаваемые изображения?
2. Каким образом достигается инвариантность результатов классификации «неизвестного» изображения к его вращению?
3. Почему изображения объектов, нормализуемые и описываемые векторами признаков согласно методике, исследованной в настоящей лабораторной работе, должны быть отделены от фона (помещены на нулевой фон)?

Литература

1. Вадим Луцив. Автоматический анализ изображений. Объектно-независимый структурный подход. – Саарбрюккен, Германия: Ламберт Академик Пабблишинг, 2011. – 300 с.
2. Луцив В.Р. Метод обработки изображений при управлении роботами в ГАП. – В кн.: ЭВМ в проектировании и производстве: Сб. статей. Вып.2 /Под общ. ред. Г.В. Орловского. – Л.: Машиностроение, Ленингр. отделение, 1985. – с.154-162.
3. Потапов А.С. Системы компьютерного зрения: реализация в библиотеке OpenCV [Текст]: учебно-методическое пособие / А. С. Потапов, М. В. Петерсон, В. А. Козлов, А. Н. Аверкин. СПб.: СПбГУ ИТМО, 2011. 83 с.
4. Bradski G. and Kaehler A. Learning OpenCV. Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo: O'Reilly Media, Inc., 2008. 556 p.
5. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. 1072 с.

ПРИЛОЖЕНИЕ

Фрагмент программы на языке C, осуществляющий ввод изображения из файла стандартного графического формата и вывод его на экран дисплея средствами библиотеки Intel OpenCV

Результаты работы написанного студентами алгоритма преобразования изображения, введенного из файла, можно визуализировать, вставив, как показано, в эту программу строки, реализующие исследуемый алгоритм, перед строками вывода изображения на дисплей.

```
#include "highgui.h"
int main( int argc, char **argv )
{
    IplImage *image = NULL;
    if( argc == 2 )
        image = cvLoadImage( argv[1], CV_LOAD_IMAGE_UNCHANGED );
    else return -1;
    if( image == NULL )
        return -2;
    /*
        Строки программной модели, реализующей исследуемый алгоритм
        обработки изображения
    */
    cvNamedWindow( "Test OpenCV", CV_WINDOW_AUTOSIZE );
    cvShowImage( "Test OpenCV", image );
    cvWaitKey(0);
    return 0;
}
```

Миссия университета – генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

КАФЕДРА КОМПЬЮТЕРНОЙ ФОТОНИКИ И ВИДЕОИНФОРМАТИКИ

Достижения в оптической науке, технике и технологиях за последние годы способствовали появлению нового направления – фотоники. Этот термин охватывает область науки и техники, связанную с использованием светового излучения (или потока фотонов) в оптических элементах, устройствах и системах.

На рубеже XX – XXI веков электронные информационные технологии достигли фундаментальных и технических пределов производительности при продолжающемся росте потребительского спроса на скорость и объем обрабатываемой и передаваемой информации. Решение данной проблемы потребовало разработки нового поколения информационно – телекоммуникационных систем, основанных на технологиях фотоники. В фотонике появилось новое динамично развивающееся направление, определяющее прогресс мировой науки и техники, – «оптоинформатика». Под «оптоинформатикой» понимают область науки и техники, связанную с исследованием, разработкой, созданием и эксплуатацией новых материалов, технологий, приборов и устройств, направленных на передачу, прием, обработку, хранение и отображение информации.

Изучение фотоники основывается на знании принципов формирования, преобразования, анализа изображений, теории построения информационных систем. Интеграция фотоники и компьютерных технологий позволяет создавать методы, которые возможно реализовать исключительно средствами компьютерной фотоники, обеспечивая развитие технологий качественно нового уровня.

По многим направлениям фотоники и оптоинформатики Россия находится на уровне промышленно – развитых стран (интегральная оптика, системы приема, обработки и отображения информации и др.), а, по некоторым – даже опережает. Приоритетными направлениями являются: волоконная оптика (работы академика Дианова Е.М. – ИОФ РАН), голография (академик

Денисюк Ю.Н. – ГОИ им. С.И. Вавилова), полупроводниковые лазеры (академик Алферов Ж.И – ФТИ РАН им. А.Ф. Иоффе), полифункциональные оптические материалы (академик Петровский Г.Т. – ГОИ им. С.И. Вавилова) и др.

Ввиду большого научного и практического значения направления "Фотоника и оптоинформатика", а также спроса на него на потребительском рынке в 2002 г. в СПбГУ ИТМО был организован факультет «Фотоники и оптоинформатики» под руководством доктора физ.-мат. наук, профессора С.А. Козлова. По инициативе профессорско-преподавательского состава, начиная с 2005 года, на факультете стала работать выпускающая кафедра «Компьютерной фотоники», которую возглавил доктор технических наук, профессор И.П. Гуров.

История кафедры началась в 1946 году. На всех этапах развития результаты научных исследований, проводимых сотрудниками кафедры, неизменно использовались в учебном процессе. Совершенствовались направления подготовки студентов, изменялось название кафедры, но всегда кафедра гордилась своими выпускниками.

Выпускники кафедры занимают видное место в оптической науке: академик РАН Ю.Н. Денисюк, изобретатель трехмерной голографии; член-корр. РАЕН, профессор Н.Г. Бахшиев, известный специалист в области спектроскопии межмолекулярных взаимодействий; Заслуженный деятель науки РФ, профессор Г.Н. Дульнев, крупный ученый в области теплофизики, долгие годы бывший ректором ЛИТМО; профессор И.М. Нагибина, исследования которой в области физической оптики получили широкое признание.

Одной из важнейших задач кафедры является организация учебного процесса и подготовка профессионалов в области компьютерной фотоники. Направление работы кафедры определяется развитием информационных технологий и компьютерных систем в области формирования, синтеза, обработки и анализа изображений на основе интеграции эффективных компьютерных систем с системами фотоники.

Проводимые исследования в области компьютерной обработки когерентных и некогерентных изображений обеспечивают решение научно-технических задач оптической томографии, цифровой голографии, синтеза, анализа, распознавания и классификации изображений.

Научным консультантом работ кафедры в области компьютерной обработки изображений – иконики – является член-корреспондент РАН М.М. Мирошников.

Кафедра проводит работы в рамках международных научных проектов в сотрудничестве с ведущими зарубежными университетами, институтами и исследовательскими лабораториями Италии, Финляндии, Франции,

Германии, Великобритании, Японии, США и других стран в области оптической когерентной томографии для биомедицинских исследований, цифровой голографии для исследования микро- и наноструктур, трехмерной фотографии микро- и макроскопических объектов, гиперспектральной обработки изображений.

В последнее время на кафедре активно развивается новое направление – видеоинформатика. Ввиду этого в 2010 году кафедра была переименована в кафедру Компьютерной фотоники и видеоинформатики. В 2011 году Университет получил статус Национального исследовательского университета, в этом есть и заслуга преподавательского коллектива кафедры Компьютерной фотоники и видеоинформатики. Эти обстоятельства позволяют обеспечивать и в дальнейшем подготовку высококлассных востребованных на рынке специалистов в области компьютерной фотоники и видеоинформатики.

Волкова Мария Анатольевна
Луцив Вадим Ростиславович

Методы обработки и распознавания изображений

Учебно-методическое пособие по лабораторному практикуму

В авторской редакции
Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Н.Ф. Гусарова

Подписано к печати

Заказ №

Тираж

Отпечатано на ризографе

Редакционно-издательский отдел
Университета ИТМО
197101, Санкт-Петербург, Кронверкский пр., 49