

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, ОСНОВАННОЕ НА РАНГОВОМ РАСПРЕДЕЛЕНИИ ИХ ПРОСТРАНСТВЕННО-ЯРКОСТНЫХ ХАРАКТЕРИСТИК

Петрик И.М.

Научный руководитель: Стоянов А.К.

Томский политехнический университет, Институт кибернетики

imp1@tpu.ru

Введение

Распознавание изображений используется во многих случаях их использования: поиск изображений по оригиналу в Интернете, цифровые архивы и библиотеки; анализ содержания космических и аэрофотоснимков, системы технического зрения и т.д. [1]. Одной из проблем, возникающих при решении этой задачи, является поиск изображений, являющихся несколько искажёнными копиями оригинала.

Ранее в [2] была показана возможность успешного распознавания искажённых копий (нечётких дубликатов) изображения. Для распознавания было применено ранговое распределение относительных мощностей кластеров их яркости.

Пиксель любого изображения имеет три независимых характеристики – яркость и две пространственные координаты. В [2] для распознавания использована только одна из характеристик – яркость, которая была кластеризована по близким значениям. Но два изображения разных объектов могут иметь очень схожие распределения яркости пикселей, и, как следствие, после кластеризации (сегментации) они будут признаны похожими. Очевидно, это приводит к ошибкам и ограничивает возможности описанного метода

Учёт значений не только яркости, но и всех характеристик должен привести к уменьшению ошибок распознавания. В данной работе исследована возможность распознавания изображений с использованием кластеризации их пространственно-яркостных характеристик.

Кластеризация числовых характеристик

Для кластеризации пространственно-яркостных характеристик изображений был выбран многошаговый метод, который был использован ранее в [2-3] для сегментации полутоновых изображений по яркости. Этот метод обладает таким преимуществом, как отсутствие необходимости предварительного задания числа кластеров.

Метод основан на использовании для кластеризации значений яркости рекуррентной нейронной сети. Такая сеть состоит из одного слоя нейронов, каждый из которых охвачен локальной обратной связью [3].

Работа отдельного нейрона сети моделируется одномерным отображением значений входного сигнала нейрона на его активационной функции

$f(x)$. В качестве такой функции мы используем сигмоиду.

Диаграмма Ламерея (рис. 1) наглядно представляет процесс такого отображения. В процессе отображения $x_{n+1} = f(x_n)$, где n – номер текущей итерации, любое значение x , попавшее в интервал между точками A и B , достигнет устойчивой точки, с заданной точностью, за одно и то же количество итераций. Именно этим свойством и обеспечивается эффект кластеризации значений яркости изображения и, как следствие, его сегментация по яркости [2-3].

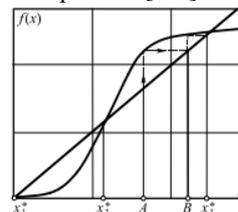


Рисунок 1 - Диаграмма Ламерея для одномерного отображения значений яркости (x_1^* и x_3^* – устойчивые точки; x_2^* – неустойчивая точка отображения)

В [2] на вход нейронов сети подаются значения яркостей пикселей изображения. Многошаговая сегментация выполняется в несколько шагов, каждый из которых использует результаты предыдущего шага в качестве входных данных. Процесс завершается после того, как энтропия изображения, полученного на очередном шаге, перестаёт меняться. Такая сегментация обеспечивает максимально упорядоченное по определённой характеристике распределение однородных областей пикселей на изображении.

В качестве характеристики изображения при распознавании в [2] используются ранговые распределения относительных мощностей выделенных кластеров его яркости. Под относительной мощностью кластера понимается доля пикселей изображения, попавших в кластер, от их общего числа. Ранговое распределение – это зависимость некоторой величины от ранга – порядкового номера в ряду её значений, упорядоченных по убыванию.

Очевидно, в нашем случае, после кластеризации пространственно-яркостных характеристик мы получим три набора кластеров: по яркости, по координатам x и y . Общий кластер должен включать в себя пиксели, одновременно отнесённые ко всем трём кластерам (яркость две координаты).

Оценка близости изображений

Для оценки близости сравниваемых изображений удобно использовать их представление в виде векторов ранговых распределений в n -мерном пространстве, где n – число рангов в распределении. Мету близости между двумя векторами изображений в n -мерном пространстве задаётся в виде угла. Пусть есть два вектора сравниваемых изображений $S = (s_1, s_2, \dots, s_n)$ и $X = (x_1, x_2, \dots, x_n)$. Тогда близость d_{sx} таких векторов определяется выражением:

$$d_{sx} = \arccos\left(\frac{\sum_i s_i * x_i}{|X| * |S|}\right) \quad (1)$$

где $|S|$ и $|X|$ в знаменателе соответственно длины векторов S и X , а в числителе стоит их скалярное произведение. Похожесть изображений S и X определяется с помощью решающего правила: $S \in X$, если $d_{sx} \leq \varepsilon$, где $\varepsilon \ll 1$ – наперёд заданная положительная величина.

Экспериментальные результаты

С целью проверки возможности применения кластеризации пространственно-яркостных характеристик для распознавания похожих изображений нами был проведён ряд экспериментов с изображениями в градациях серого цвета. Для сегментации использована модификация нейронной сети, описанная в [2-3].

Искажённые копии изображения создавались гауссовым размытием яркостей пикселей оригинала, выполненным с помощью растрового редактора Adobe Photoshop. На рис. 2 приведены некоторые из использованных для распознавания изображений.



Рисунок 2 - Изображения лица мужчины: а) оригинал; б) искажённое гауссовым размытием с радиусом 6 пикселей); в) искажённое гауссовым размытием с радиусом 8 пикселей

На рис. 3 приведены ранговые распределения мощностей кластеров яркости и координат для оригинала изображения и размытых копий. По ординате откладывались относительные значения мощности кластеров, по абсциссе откладывались номера рангов кластеров.

Форма кривых на рис. 3 позволяет качественно оценить похожесть соответствующих им изображений. Видно, что хорошее совпадение распределений оригинала изображения и его копий наблюдается для значений радиусов размытия, примерно, до 6 пикселей

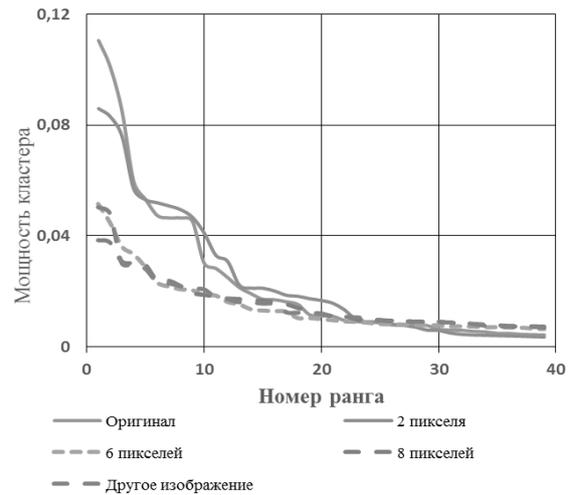


Рисунок 3 - Ранговое распределение мощностей кластера в различных изображениях

Результаты количественного сравнения близости ранговых распределений мощностей кластеров для оригинала, дубликатов и других изображений приведены в таблице. Близость оценивалась по косинусному расстоянию d_{sx} , вычисленному по формуле (1). Пороговое расстояние, при котором изображение считается совпадающим с оригинальным, было выбрано равным $d_0 \leq 0,157$ [2]. Из таблицы видно, что количественная оценка близости подтверждает вывод, полученный при качественном изучении графика.

Радиус искажения	Косинусное расстояние
Оригинал	0
2 пикселя	0,015
4 пикселя	0,072
6 пикселей	0,36
Другое изображение (пейзаж)	0,37

Выводы

1. Кластеризация пространственно-яркостных характеристик пикселей позволяет выполнять распознавание изображения.
2. В экспериментах при заданном пороге сравнения были определены как похожие изображения, искажённые гауссовым размытием радиусом до 6 пикселей.

Литература

1. Л. Шапиро, Дж. Стокман. Компьютерное зрение. – М.: Бином. Лаборатория знаний, 2006. – 752 с.
2. В.Б. Немировский, А.К. Стоянов. Распознавание нечётких дубликатов изображений, основанное на ранговом распределении мощностей кластеров яркости / Компьютерная оптика, 2014, том 38, №4. – С. 814-817
3. В.Б. Немировский, А.К. Стоянов. Сегментация изображений с помощью рекуррентной нейронной сети / Известия Томского политехнического университета. – 2012. – Т. 321, № 5. – С. 205-210.