электронный журнал
МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". Эл No. ФС77-51038.

 # 05, май 2016

УДК 004.93'12

Классификация изображений на основе дескрипторов цвета, текстуры и границ

Крайникова А.В., *студент*
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

Научный руководитель: Лычков И.И., старший преподаватель
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»
iu3@bmstu.ru

Введение

В связи с быстрым ростом объемов мультимедийной информации возникают задачи автоматического поиска интересующей информации и навигации по ней. Одним из подходов к поиску изображений по содержанию является классификация изображений, то есть отнесение изображения-запроса к одному из заранее определенных классов изображений и выдача примеров изображений данного класса. В настоящее время задаче классификации изображений уделяется большое внимание в научной литературе [1-9].

Классификация изображений основывается на анализе вычисляемых характеристик изображений на уровне пикселей. В связи с этим возникает «семантический разрыв», когда человек и машина сравнивают изображения по-разному: человек сравнивает смысловую наполненность, а машина – низкоуровневые признаки. Проблема семантического разрыва до сих пор не решена. Для классификации изображений могут использоваться глобальные дескрипторы, описывающие изображение целиком, или локальные дескрипторы, описывающие значимые элементы изображения.

Целью настоящего исследования является разработка и реализация системы классификации изображений по общему подобию, то есть по глобальным дескрипторам. Классификация изображений является актуальной задачей, возникающей при автоматическом анализе мультимедийных баз данных большого размера. Например, для поиска изображений, похожих на заданное изображение-запрос, или для обратного поиска, когда для заданного изображения-запроса требуется найти информацию об объекте, запечатленном на этом изображении.

Решению задачи классификации изображений посвящено большое количество исследований. В работах [1-9] рассматриваются основные направления исследований и анализ состояния предметной области.

В настоящей работе в качестве низкоуровневых признаков изображения выбраны цвет, границы и текстура, представленные с помощью дескрипторов стандарта MPEG-7, используются различные комбинации признаков. Классификация изображений выполняется двумя методами: с помощью классификатора k ближайших соседей и с помощью самоорганизующейся карты Кохонена. Для проведения эксперимента используется специально подготовленная база изображений ландшафтов.

Далее статья структурирована следующим образом. В разделе 1 рассматриваются выбранные дескрипторы цвета, границ и текстуры, а также результаты визуализации дескрипторов с помощью специально разработанной программы. В разделе 2 описывается предложенная система классификации изображений. В разделе 3 приводится описание проведенного эксперимента и обсуждение его результатов. Статья завершается заключением.

1. Дескрипторы

MPEG-7 [10] - это стандарт, который предлагает полный набор средств описания изображений и видеозаписей. Данный стандарт представляет методы извлечения числовых признаков из мультимедийного материала, что важно быстрой фильтрации и поиска нужных материалов. Главными элементами стандарта являются: дескрипторы, схемы описания, язык описания определений и системные средства.

В настоящей работе для классификации изображений использовались три вида дескрипторов: цвета, текстуры и границ. Для вычисления дескрипторов использовалась библиотека MPEG7FexLib [11]. Элементы дескрипторов выводятся в файл текстового формата CSV в виде матрицы, в которой столбцы соответствуют элементам дескрипторов для отдельного изображения, а строки соответствуют различным изображениям. Чтобы эффективно использовать дескрипторы для статистического анализа, элементы дескрипторов приводятся к одному диапазону значений с помощью нормализации [12] согласно формуле (1):

$$x'_{ik} = \frac{x_{ik} - min_k}{max_k - min_k}, \quad (1)$$

где min_k и max_k – минимальное и максимальное значение в столбце k. Результирующее значение x'_{ik} лежит в диапазоне [0, 1]. При такой нормализации форма распределения значений в строках и столбцах сохраняется неизменной.

Ниже описываются выбранные дескрипторы, их формат представления в стандарте MPEG-7, а также разработанная программа визуализации для каждого дескриптора.

Цвет является одним из наиболее широко используемых визуальных характеристик изображений и является достаточно надежным для классификации изображений. В настоящей работе был выбран дескриптор доминантного цвета**.** Дескриптор состоит из процентного содержания цвета и значений компонентов цветового пространства. На рисунке 1 показаны примеры выполнения программы, разработанной для визуализации дескриптора доминантного цвета.

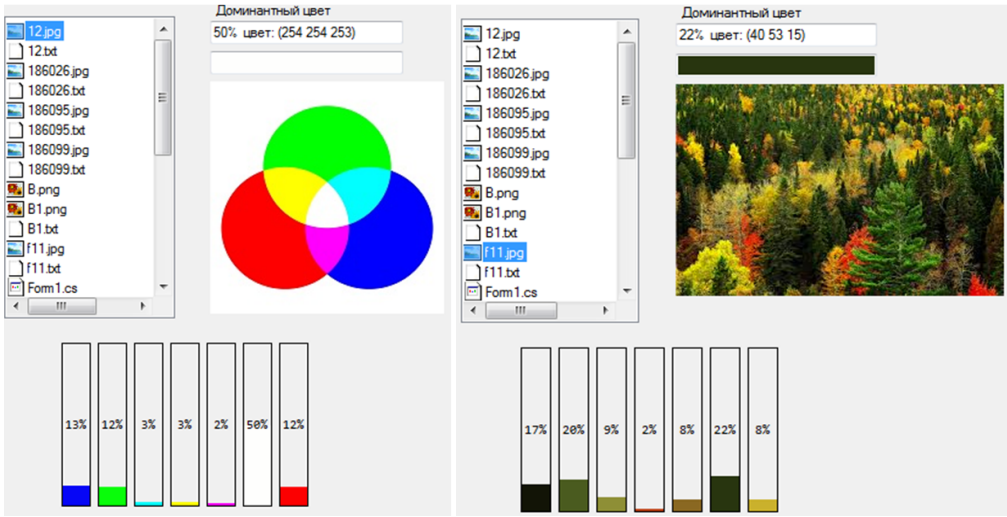


Рис. 1. Визуализация дескрипторов доминантного цвета

Дескриптор доминантного цвета в стандарте MPEG-7 использует цветовую модель RGB, т.е. представляет цвета изображения в виде комбинации трех основных цветов: красного, зеленого и синего. Для выделения цветового тона предлагается преобразовать цветовую модель RGB в цветовую модель HSV [13], где Hue — цветовой тон, Saturation — насыщенность, Value — яркость.

Пространственное распределение границ на изображении также является информативной характеристикой изображения. Для представления границ в настоящей работе был выбран дескриптор гистограммы границ. Он представляет собой гистограмму, в которой границы подразделяются на 5 типов: вертикальные, горизонтальные, диагональ 45º, диагональ 135º и ненаправленные границы. На рисунках 2 (а, б) показаны примеры выполнения программы, разработанной для визуализации дескриптора границ. На гистограммах первый столбец соответствует вертикальным линиям, второй – горизонтальным, а третий, четвертый, пятый соответствуют 45º, 135º и ненаправленным границам.

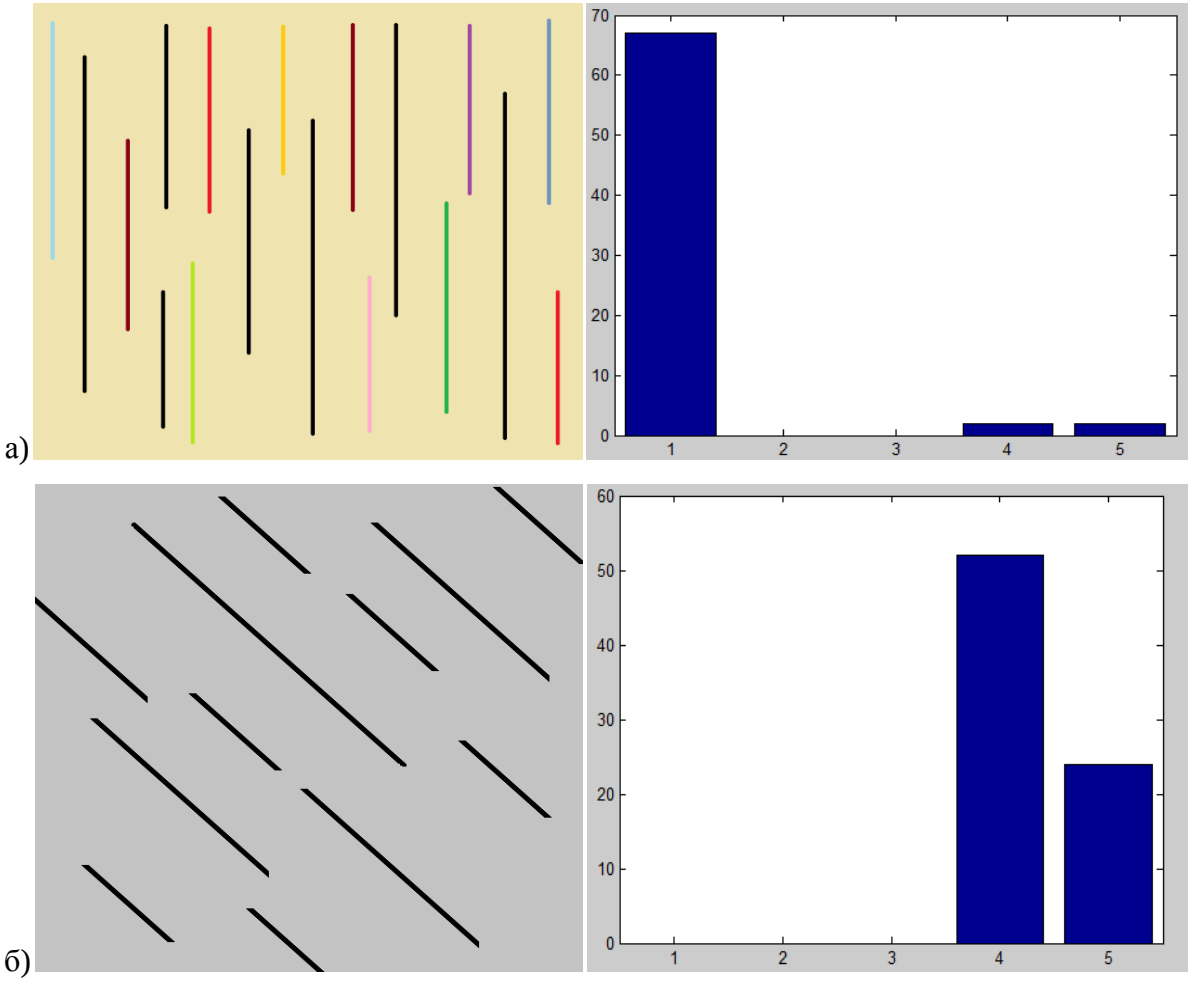


Рис. 2. Визуализация дескрипторов границ (слева – входное изображение, справа – соответствующая гистограмма)

Для представления текстуры изображения был выбран дескриптор однородности текстуры. Дескриптор однородности текстуры характеризует текстуру изображения с помощью средней энергии изменения яркости изображения и отклонений энергии изменения яркости изображения от средней энергии для набора частотных диапазонов.

На рисунке 4 изображены гистограммы суммарной энергии по каждому из 6 направлений и по каждому из 5 частотных диапазонов для тестового изображения, приведенного на рисунке 3. Для данного тестового изображения гистограммы близки к равномерным, что позволяет сделать вывод об однородной текстуре изображения.

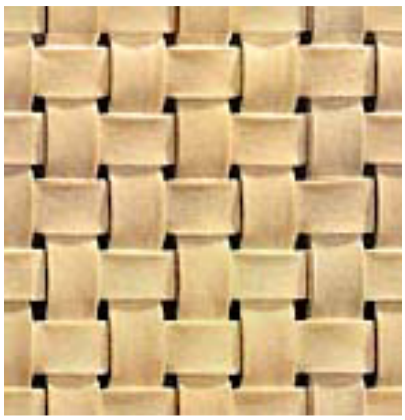


Рис. 3.Тестовое изображение.

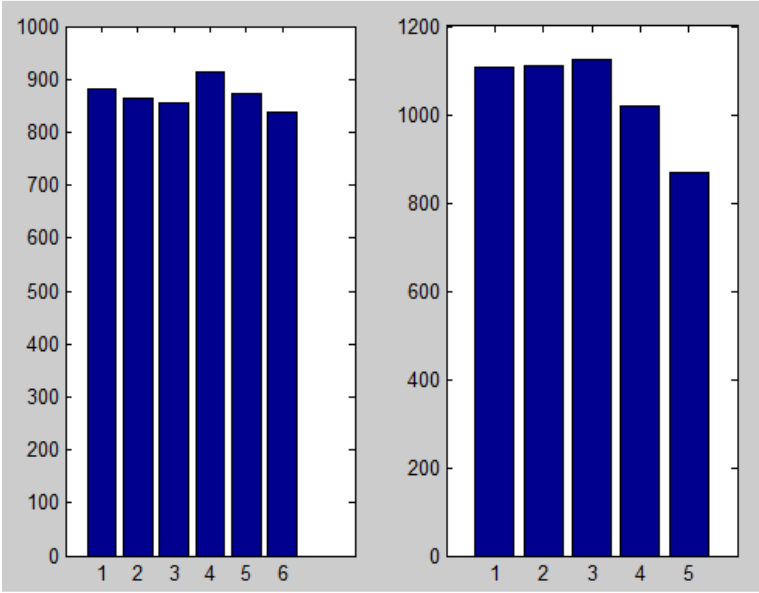


Рис. 4. Гистограммы суммарной энергии по каждому направлению
и по каждому частотному диапазону

2. Система классификации изображений

В декабре 2006 года на международной конференции по интеллектуальному анализу данных в Гонконге [14] были выявлены наиболее важные алгоритмы интеллектуального анализа данных в исследовательском сообществе. Эти алгоритмы охватывают классификацию, кластеризацию, статистическое изучение, анализ взаимосвязей, которые являются важными темами для интеллектуального анализа данных, исследования и разработки. На конференции были представлены такие алгоритмы как: алгоритм k-means, машина опорных векторов (SVM), EM-алгоритм, AdaBoost, метод Naive Bayes и другие.

Для классификации изображений выберем два метода, один из которых относится к обучению «без учителя», другой – «с учителем». При обучении «с учителем» метки принадлежности обучающих изображений к классам задаются исследователем вручную. При обучении «без учителя» метки принадлежности обучающих изображений к классам назначаются автоматически так, чтобы внутри каждого кластера изображения были максимально сходны, а между разными кластерами – максимально различны.

В качестве методов классификации выберем метод k ближайших соседей и метод самоорганизующихся карт Кохонена.

Метод k-ближайших соседей [15] реализует обучение «с учителем», его преимуществами являются устойчивость к аномальным выбросам, относительно простая программная реализация и легкость интерпретации результата работы алгоритма человеком.

Самоорганизующиеся карты Кохонена [16] относятся к моделям нейронных сетей с обучением «без учителя» и удобны для визуального представления многомерных свойств объектов на двумерной карте. Самоорганизующаяся карта состоит из ячеек, количество которых задается исследователем. Изображения с близкими векторами признаков попадают в одну ячейку или в соседние ячейки. Преимуществами данного метода являются устойчивость к зашумленным данным, быстрое обучение, а также возможность визуализации многомерных входных данных. Недостатком метода является зависимость результата обучения от начальных параметров сети.

В зависимости от выбранного метода обучение классификатора будет производиться «с учителем» или «без учителя». На рисунках 5 (а, б) представлены блок-схемы классификации изображений для двух методов.

3. Описание эксперимента

Предложенная система классификации изображений была реализована программно на платформе Matlab и испытана экспериментально.

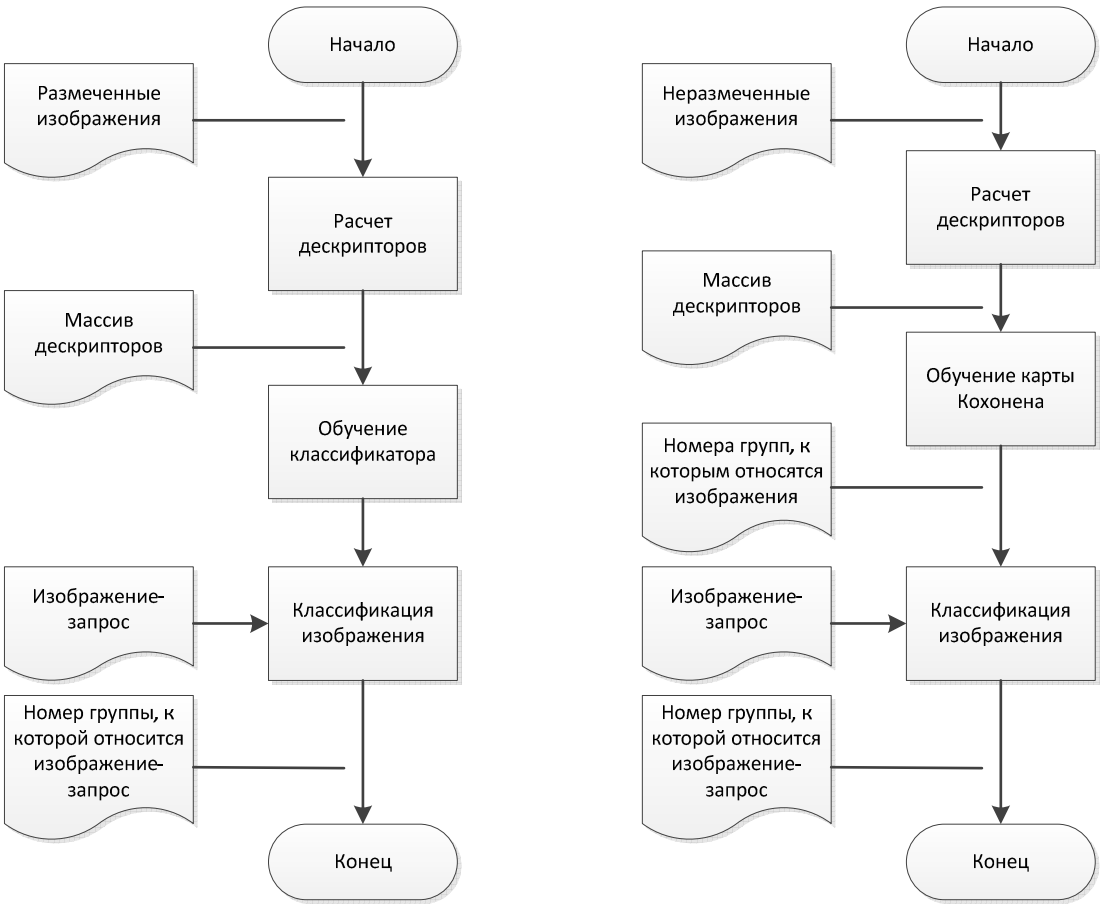
С помощью поисковой системы Google была подобрана база из 120 изображений по четырем категориям: лес, море, горы и здания. Изображения были приведены к разрешению около 50 тысяч пикселей. На рисунке 6 приведены некоторые изображения из этой базы.

Для метода самоорганизующихся карт Кохонена было произведено обучение карты с использованием трех различных комбинаций признаков:

– Только по доминантному цвету;

– По доминантному цвету и гистограмме границ;

– По доминантному цвету, гистограмме границ и однородности текстуры.



(а) (б)

Рис. 5. Схемы обработки изображений для двух методов:
а – классификатор k ближайших соседей; б – карта Кохонена

Классификация изображений методом самоорганизующихся карт Кохонена по нормализованным дескрипторам производилась с помощью библиотеки Neural Network Toolbox из пакета Matlab. Поскольку база изображений содержала 4 категории, размер карты был установлен равным 2×2, количество итераций обучения было установлено равным 200, значения остальных параметров были установлены по умолчанию.

При обучении карты происходит кластеризация обучающих изображений, то есть их автоматическое распределение по ячейкам карты. Каждой ячейке присваивался класс изображений, попавших в данную ячейку в наибольшем количестве. Изображения данного класса, попавшие в данную ячейку, считались верно кластеризованными. Остальные изображения, попавшие в данную ячейку, считались неверно кластеризованными. Точность кластеризации вычислялась как отношение количества

верно кластеризованных изображений по всем ячейкам к общему количеству обучающих изображений.

Рис. 6. Изображения из обучающей базы

При использовании одного дескриптора доминантного цвета наибольшая точность кластеризации получилась для изображений леса и составила 85%, наименьшая точность кластеризации получилась для изображений гор и составила 26%. Средняя точность кластеризации составила 50%.

При использовании дескрипторов доминантного цвета и гистограммы границ наибольшая точность кластеризации получилась для изображений моря и составила 82%, наименьшая точность кластеризации получилась для изображений гор и составила 48%. Это можно объяснить нечеткими линиями наклона гор под 45 и 135 градусов, а также наличием на некоторых изображениях зелени и неба. Средняя точность кластеризации составила 70%.

На рисунке 7 изображено распределение изображений по кластерам (слева) и положение центров кластеров (справа) при использовании двух дескрипторов. Для анализа точности кластеризации была вычислена нормированная матрица беспорядка, представленная в таблице 1. Значения точности кластеризации по каждому классу изображений указаны на главной диагонали матрицы.

После добавления третьего дескриптора однородности текстуры средняя точность кластеризации составила 53%, то есть уменьшилась по сравнению с обучением по двум дескрипторам. Точность кластеризации изображений гор составила 33%, однако точность кластеризации леса немного улучшилась и составила 85%.

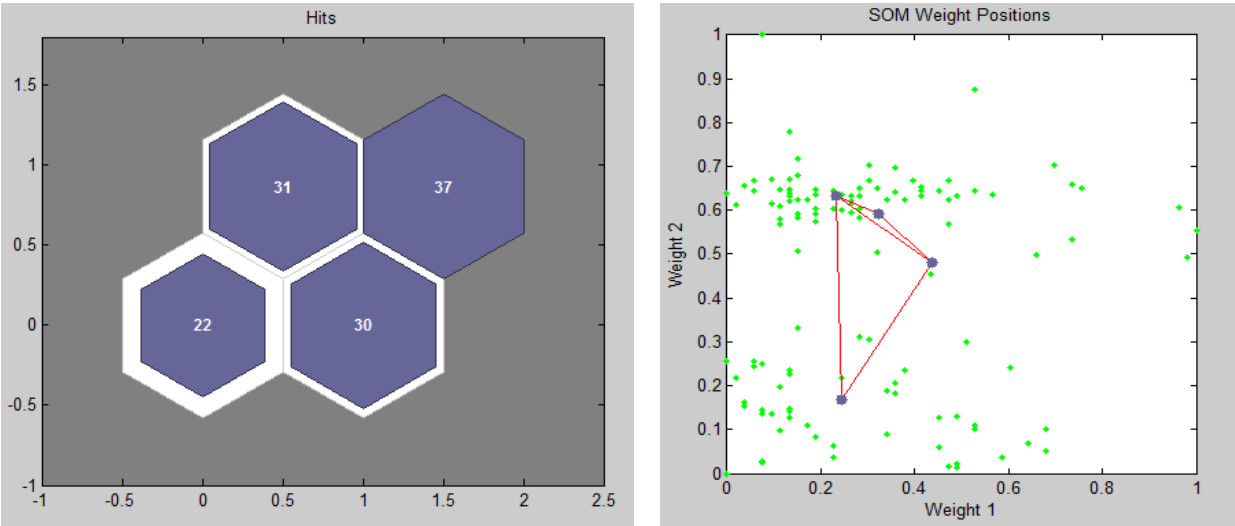


Рис. 7. Кластеризация для двух нормализованных дескрипторов

Таблица 1

Нормированная матрица для двух дескрипторов

		Предсказанный класс			
		Здание	Лес	Горы	Море
Действительный класс	Здание	**0,76667**	0,027027	0,16129	0,045455
	Лес	0,033333	**0,7027**	0,096774	0
	Горы	0,166667	0,189189	**0,48387**	0,136364
	Море	0,033333	0,081081	0,258065	**0,81818**

Оценка точности классификации проводилась на 20 тестовых изображениях методом скользящего контроля, который предполагает исключение из базы одного

изображения в качестве тестового и обучение классификатора на оставшихся изображениях. Для метода k ближайших соседей параметр k был принят равным 5.

Точность классификации вычислялась следующим образом:

$$\text{Точность} = \frac{\text{Количество верно классифицированных изображений}}{\text{Количество тестовых изображений}} * 100\ \% \qquad (2)$$

При использовании метода k ближайших соседей были верно классифицированы 13 изображений из 20, и точность классификации составила 65% (таблица 2). При использовании метода карт Кохонена точность классификации составила 50% (таблица 3).

Таблица 2

Классификация 20 изображений методом k ближайших соседей

		Предсказанный класс			
		Здание	Лес	Горы	Море
Действительный класс	Здание	**3**	0	2	0
	Лес	2	**3**	0	0
	Горы	0	0	**3**	2
	Море	0	0	1	**4**

Таблица 3

Классификация 20 изображений методом карт Кохонена

		Предсказанный класс			
		Здание	Лес	Горы	Море
Действительный класс	Здание	**2**	3	0	0
	Лес	0	**5**	0	0
	Горы	0	4	**1**	0
	Море	0	1	2	**2**

Заключение

В рамках проведенных исследований была разработана система классификации изображений на основе глобальных дескрипторов цвета, текстуры и границ. Была произведена оценка точности разработанной системы с использованием различных

комбинаций дескрипторов изображения на специально подготовленной базе изображений ландшафтов. Наибольшая точность классификации была достигнута с использованием комбинации из двух дескрипторов: доминантного цвета и гистограммы границ.

Для классификации изображений использовались два метода: метод k ближайших соседей и метод самоорганизующихся карт Кохонена. Метод k-ближайших соседей имеет более высокие показатели точности, поскольку при обучении он использует изображения, размеченные человеком. Метод самоорганизующихся карт Кохонена позволяет проводить обучение на огромных мультимедийных базах изображений, которые невозможно разметить вручную. В эксперименте точность классификации изображений обоими методами оказалась недостаточно высокой вследствие ограничений применимости глобальных дескрипторов изображений.

В дальнейшем планируется повысить точность классификации изображений за счет сочетания глобальных дескрипторов с локальными дескрипторами участков изображения.

Список литературы

[1]. Десятников И.Е. Поиск изображений по визуальному содержанию в графических базах данных и сети Интернет // Информационные технологии и вычислительные системы. 2013. № 2. С. 88-95.

[2]. Boomilingam T., Subramaniam M. Review on CBIR Trends and Techniques to Upgrade Image Retrieval // International Review on Computers and Software (IRECOS). 2014. T. 9. № 7. P. 1227-1240.

[3]. Murugesan S., Ansong E. D., Thirunavukarasu N. Fast Query Point Movement Techniques with Efficient Feature Extraction for CBIR Systems // Journal of Educational Policy and Entrepreneurial Research. 2015. T. 2. № 3. P. 118-125.

[4]. Yang J., Zhu S. Narrowing semantic gap in content-based image retrieval // 2012 International Conference on Computer Distributed Control and Intelligent Enviromental Monitoring. IEEE, 2012. P. 433-438.

[5]. Venkatakrishnan S. Information Theoretic Methodology For Retrieving Similar GIS Rasters : дис. University of Cincinnati, 2014. P. 101.

[6]. Veganzones M. A., Graña M. A spectral/spatial CBIR system for hyperspectral images // Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of. 2012. T. 5. № 2. P. 488-500.

[7]. Gupta E., Kushwah R. S. Combination of global and local features using DWT with SVM for CBIR // Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 2015 4th International Conference on. IEEE, 2015. P. 1-6.

[8]. Hua K. A. et al. Saliency-based CBIR system for exploring lunar surface imagery // Proceedings of the 121th. ASEE annual conference and exhibition. 2014. P. 1-18.

[9]. Milovanovic M., Minovic M., Starcevic D. Walking in colors: human gait recognition using kinect and CBIR // MultiMedia, IEEE. 2013. T. 20. № 4. P. 28-36.

[10]. Manjunath B. S., Salembier P., Sikora T. Introduction to MPEG-7: multimedia content description interface. John Wiley & Sons, 2002. T. 1. P. 371.

[11]. Software. Avaliable at: http://www.cs.bilkent.edu.tr/~bilmdg/bilvideo-7/Software.html, (accessed 25.02.2015).

[12]. Eidenberger H. How good are the visual MPEG-7 features? // Visual Communications and Image Processing 2003. International Society for Optics and Photonics, 2003. P. 476-488.

[13]. Sural S., Qian G., Pramanik S. Segmentation and histogram generation using the HSV color space for image retrieval // Image Processing. 2002. Proceedings. 2002 International Conference on. IEEE, 2002. T. 2. P. II589-592.

[14]. Wu X. et al. Top 10 algorithms in data mining // Knowledge and information systems. 2008. T. 14. № 1. P. 1-37.

[15]. Sutton O. Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction // University lectures, University of Leicester. 2012. P. 1-10.

[16]. Kohonen T. The self-organizing map // Proceedings of the IEEE. 1990. T. 78. №. 9. P. 1464-1480.