

НЕЧЕТКИЕ СИСТЕМЫ И МЯГКИЕ ВЫЧИСЛЕНИЯ

НАУЧНЫЙ ЖУРНАЛ

Российской ассоциации нечетких систем и мягких вычислений

Том 6, номер 1, июнь 2011

Основан в 2006 году. Выходит 2 раза в год.

Учредитель: ФГБОУ ВПО «Тверской государственный университет»

Главный редактор:

170002, Садовый переулок, 35

Тел./факс: (4822) 58-54-10

E-mail: Alexander.Yaznenin@tversu.ru

Ответственный секретарь:

170002, Садовый переулок, 35

Тел./факс: (4822) 58-57-43

E-mail: soldis@tversu.ru

МЕЖДУНАРОДНАЯ РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Почетный редактор проф. Л. Заде (Калифорнийский университет, Беркли, США),
гл. редактор д.ф.-м.н., проф. Язенин А.В. (Тверской госуниверситет, Тверь, Россия),
зам. гл. редактора д.ф.-м.н., проф. Батыршин И.З. (Мексиканский нефтяной институт, Мехико, Мексика),
отв. секретарь к.ф.-м.н. Солдатенко И.С. (Тверской госуниверситет, Тверь, Россия),
д.ф.-м.н., проф. Бенинг В.Е. (МГУ, Москва, Россия),
д.т.н., проф. Берштейн Л.С. (Южный Федеральный университет, Таганрог, Россия),
д.т.н., проф. Борисов А.Н. (Рижский технический университет, Рига, Латвия),
dr.habil. Вагенкнехт М. (Университет прикладных наук, Циттау, Германия),
академик РАН, д.ф.-м.н., проф. Евтушенко Ю.Г. (ВЦ РАН, Москва, Россия),
чл.-корр. РАН, д.т.н., проф. Каляев И.А. (Южный Федеральный университет, Таганрог, Россия),
д.ф.-м.н., проф. Каркищенко А.Н. (Южный Федеральный университет, Таганрог, Россия),
академик Польской академии наук, проф. Каспшик Я. (Институт системных исследований Польской академии наук, Варшава, Польша),
д.т.н., проф. Курейчик В.М. (Южный Федеральный университет, Таганрог, Россия),
д.т.н., проф. Поспелов Д.А. (ВЦ РАН, Москва, Россия),
д.ф.-м.н., проф. Редько В.Г. (Институт оптико-нейронных технологий РАН, Москва, Россия),
д.т.н., проф. Соколов А.Ю. (Национальный аэрокосмический университет, Харьков, Украина),
д.ф.-м.н., проф. Ульянов С.В. (Ямаха, Япония),
чл.-корр. РАН, д.ф.-м.н., проф. Флеров Ю.А. (ВЦ РАН, Москва, Россия),
д.ф.-м.н., проф. Хохлов Ю.С. (Тверской госуниверситет, Тверь, Россия),
д.ф.-м.н., проф. Цурков В.И. (ВЦ РАН, Москва, Россия),
д.т.н., проф. Ярушкина Н.Г. (Ульяновский технический университет, Ульяновск, Россия).

FUZZY SYSTEMS AND SOFT COMPUTING

SCIENTIFIC JOURNAL
of Russian Association for Fuzzy Systems and Soft Computing

Volume 6, number 1, June 2011

Founded in 2006. Published 2 times per year.

Founder: Tver State University

Editor-in-chief:

170002, Russia, Tver, Sadoviy, 35
Phone/fax: 7 (4822) 58-54-10
E-mail: Alexander.Yazenin@tversu.ru

Responsible Assistant:

170002, Russia, Tver, Sadoviy, 35
Phone/fax: 7 (4822) 58-57-43
E-mail: soldis@tversu.ru

INTERNATIONAL EDITORIAL BOARD

*Honorary editor prof. Zadeh L.A. (University of California, Berkeley, USA),
Editor-in-chief Doctor of Science, prof. Yazenin A.V. (Tver State University, Tver, Russia),
Deputy editor-in-chief Doctor of Science, prof. Batyrshin I.Z. (Mexican Petroleum Institute, Mexico),
Responsible assistant Soldatenko I.S. (Tver State University, Tver, Russia),*

Doctor of Science, prof. Bening V.E. (Moscow State University, Moscow, Russia),
Doctor of Science, prof. Berstein L.S. (Southern Federal University, Taganrog, Russia),
Doctor of Science, prof. Borisov A.N. (Technical University of Riga, Riga, Latvia),
dr.habil. Wagenknecht M. (University of Applied Sciences, Zittau, Germany),
academician of RAS, Doctor of Science, prof. Evtushenko Y.G. (CC RAS, Moscow, Russia),
corresponding member of RAS, Doctor of Science, prof. Kalyaev I.A. (Southern Federal University, Taganrog, Russia),
Doctor of Science, prof. Karkishenko A.N. (Southern Federal University, Russia),
Member of Polish Academy of Science, prof. Kacprzyk J. (System Research Institute, Warsaw, Poland),
Doctor of Science, prof. Kureichik V.M. (Southern Federal University, Taganrog, Russia),
Doctor of Science, prof. Pospelov D.A. (Computing Center RAS, Moscow, Russia),
Doctor of Science, prof. Red'ko V.G. (Optical neuronal technologies institute of RAS, Moscow, Russia),
Doctor of Science, prof. Sokolov A.Y. (National Aerospace University, Kharkov, Ukraine),
Doctor of Science, prof. Ulyanov S.V. (Yamaha, Japan),
corresponding member of RAS, Doctor of Science, prof. Flerov Y.A. (Computing Center RAS, Moscow, Russia),
Doctor of Science, prof. Khokhlov Y.S. (Tver State University, Tver, Russia),
Doctor of Science, prof. Tsurkov V.I. (Computing Center RAS, Moscow, Russia),
Doctor of Science, prof. Yarushkina N.G. (Ulyanovsk Technical University, Ulyanovsk, Russia).

СОДЕРЖАНИЕ

Юбилиары

К 90-летию Лотфи Заде	5
<i>Lotfi Zadeh</i>	
Factual Information about the Impact of Fuzzy Logic	11

Статьи

<i>Молодцов Д.А.</i>	
Мягкие множества и прогнозирование	15
<i>Броневич А.Г., Мельниченко А.С.</i>	
Автоматическая аннотация изображений на основе классификации статистических классов.....	31
<i>Берштейн Л.С., Боженюк А.В.</i>	
Нечеткие инварианты нечетких графов и гиперграфов	43
<i>Филатова Н.Н., Миловидов А.</i>	
Анализ и интерпретация аудиограмм в пространстве нечетких признаков	55
<i>Ульянов С.В., Литвинцева Л.В., Мишин А.А., Сорокин С.В., Фукуда Т., Тялюшкина О.Ю., Колбенко Е.В., Нефедов Н.Ю., Петров С.П., Полунин А.С.</i>	
«Парадокс» квантовой самоорганизации баз знаний и робастные интеллектуальные системы управления	67
Правила для авторов	106

УДК 004.932.72'1

АВТОМАТИЧЕСКАЯ АННОТАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ КЛАССИФИКАЦИИ СТАТИСТИЧЕСКИХ КЛАССОВ¹

Броневич А.Г.* , Мельниченко А.С.**

*ОАО «Научно-исследовательский и проектно-конструкторский институт
информатизации, автоматизации и связи
на железнодорожном транспорте», г. Москва
**Таганрогский технологический университет
Южного федерального университета, г. Таганрог

Поступила в редакцию 19.07.2011, после переработки 18.08.2011.

В данной работе рассматривается проблема автоматического аннотирования изображений набором ключевых слов, что позволяет осуществлять поиск изображений в больших коллекциях по текстовому запросу. Рассматривается общая схема аннотации с использованием глобальных низкоуровневых признаков изображений, представляемых как статистические классы. С помощью процедуры классификации статистических классов, основанной на предлагаемой мере включения, производится построение вторичных информативных признаков изображений, по которым и производится классификация изображений по ключевым словам.

This work is devoted to the problem of automatic image annotation. This problem consists in assigning words of a natural language to an arbitrary image by analyzing textural characteristics (low-level features) of images without any other additional information. It can help to extract intellectual information from images and to organize searching procedures in a huge image base according to a textual query. We propose the general annotation scheme based on the statistical classes and their classification. This scheme consists in the following. First we derive the low-level features of images that can be presented by histograms. After that we represent these histograms by statistical classes and compute secondary features based on introduced inclusion measures of statistical classes. The automatic annotation is produced by aggregating secondary features using linear decision functions.

Ключевые слова: автоматическая аннотация изображений, поиск изображений, низкоуровневые признаки, классификация, статистические классы, меры включения.

Keywords: automatic image annotation, image retrieval, low-level features, statistical classes, inclusion measures.

¹Работа выполнена при поддержке РФФИ, проекты: №10-07-00135, №10-07-00478, №11-07-00591.

Введение

В настоящее время широкое распространение получили базы данных изображений размером в тысячи и миллионы, а количество изображений в Интернете почти не поддается подсчету. Поиск изображений в таких больших коллекциях наиболее удобным для пользователя способом с помощью текстового запроса затруднен вследствие отсутствия связанных с изображениями ключевых слов. Наиболее перспективным способом решения этой проблемы является использование методов автоматической аннотации изображений, позволяющих построить для изображения набор слов, описывающих его визуальное содержание.

В настоящее время существует несколько подходов к решению задачи автоматической аннотации, отличающихся главным образом способом построения признаков изображений и типом используемого классификатора [1, 2]. Так, в качестве признаков изображения может использоваться региональное представление, когда каждый отдельный регион (область) изображения описывается своим вектором признаков, либо глобальное представление, когда все изображение описывается одним вектором низкоуровневых признаков. Для построения региональных представлений изображений существует две стратегии: деление изображения на определенное количество блоков фиксированного размера или на некоторое количество регионов произвольной формы (полученных с помощью алгоритмов кластеризации). Хотя разбиение изображения на блоки проще реализуемо, однако такое представление мало согласуется с семантическим разбиением изображения на объекты. Однако достаточно аккуратная автоматическая сегментация изображений является в свою очередь сложной проблемой.

Для вычисления как глобальных, так и региональных представлений могут использоваться цветовые и текстурные характеристики, характеристики формы регионов. Так как вычисление признаков изображения представляет собой, как правило, самый вычислительно сложный этап алгоритма аннотации, процедура их вычисления должна быть одновременно достаточно простой и в то же время давать достаточное для дальнейшей работы количество информации. В большинстве работ, использующих какое-либо региональное представление, используется комбинация из цветовых и текстурных признаков: гистограмм цветов и распределений градиентов [2], средний цвет, моменты инерции, цветовые и текстурные вариации [3], гистограммы в различных цветовых пространствах, коэффициенты вейвлет-разложений [4].

Среди подходов к построению модели аннотирования (классификатора, решающих функций) большую группу составляют вероятностные классификаторы. Так, байесовский классификатор может быть использован для описания зависимости между случайными переменными (признаками) и построения аппроксимации плотности из совместного распределения. Существует группа методов на основе байесовской классификации, использующих понятия релевантности и «перевода» признаков в слова, заимствованные из области информационного поиска и машинного лингвистического перевода. В рамках первых моделей из этой группы [2, 3] изображение с помощью регулярной сетки делится на равные прямоугольные части, каждый вектор признаков отдельного региона соотносится с одним ключевым словом с помощью оценки вероятностей одновременного появления на данном изображении ключевого слова и полученного признака. Усовершенствованием этих моделей является модель кросс-медиа поиска по релевантности (Cross-Media Relevance Model, CMRM) [7, 8]. Строится вероятностная модель, которая позволяет предсказать на основе набора регионов нового изобра-

жения вероятность соответствия ему ключевого слова. Главное отличие метода от предыдущих состоит в том, что слова присваиваются всему изображению в целом, а не отдельным регионам, что лучше соответствует модели человеческого восприятия изображений. Кроме того, в [6] сделана попытка использования CMRM модели совместно с глобальными признаками изображений, а также учета корреляции слов, появляющихся в аннотациях изображений обучающей выборки.

Другой группой методов являются методы, основанные на кластеризации пространства признаков с использованием некоторой специальной метрики и классификатора. В [5] производится кластеризация сингнатур с использованием алгоритма кластеризации дискретных распределений, основанного на расчете расстояния между сингнатурами, представляющего сумму квадратов расстояний Маллоу (Mallows distance) между парами векторов признаков. Вычисление этого расстояния сводится к решению транспортной задачи, для чего существуют эффективные алгоритмы. Предложенный в данной работе алгоритм реализован в одной из реально действующих в Интернет систем автоматического аннотирования и поиска изображений Alipr [9].

Таким образом, мы видим, что автоматическую аннотацию изображений можно рассматривать как проблему классификации: предъявляемому изображению необходимо поставить в соответствие список ключевых слов из заданного словаря. Поскольку автоматическое получение сложных описаний, отражающих семантическую зависимость объектов на изображении, крайне затруднительно для широкого класса изображений, то далее мы будем рассматривать наиболее простую модель, представляющую описание изображения с помощью списка ключевых слов, упорядоченных по релевантности их изображению. В качестве слов для аннотации изображений будут рассматриваться слова общего плана такие, как «море», «дерево», «ландшафт», «город», характеризующие изображение в целом. Очевидно, что такие характеристики изображения могут быть выявлены с помощью низкоуровневых признаков изображений, таких как цвет, текстура, градиент. Ясно, что установить явную связь между низкоуровневыми признаками изображений и ключевыми словами чрезвычайно затруднительно. Кроме того, для качественной классификации требуется извлечь из изображения как можно больше признаков, что приводит к высокой размерности образов. Эти особенности рассматриваемой задачи приводят к следующей общей структуре системы автоматической аннотации, изображенной на рис. 1. Согласно этой схеме, из изображения вначале выделяются низкоуровневые признаки, затем на основе этих признаков строятся вторичные признаки, которые можно рассматривать как результат действия некоторого алгоритма снижения размерности исходных образов. После этого по вторичным признакам строится аннотация, основанная на использовании решающих функций. Отметим, что параметры алгоритмов выделения вторичных информативных признаков и решающие функции находятся на этапе обучения, во время которого анализируются обучающие выборки аннотированных изображений.

1. Выделение низкоуровневых признаков

Выделение низкоуровневых признаков является важным этапом работы алгоритма аннотации, так как составляет основу построения представления изображений. Используемые для задачи автоматической аннотации низкоуровневые

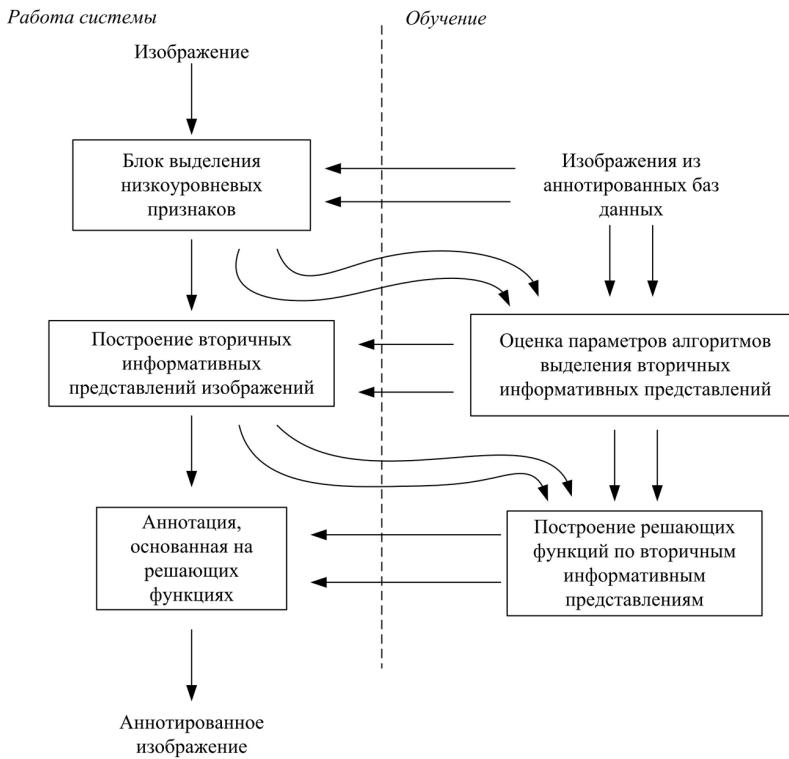


Рис. 1: Общая схема автоматической аннотации изображений

признаки должны отвечать визуальному восприятию человеком сходства между изображениями и удовлетворять некоторым свойствам, позволяющим рассматривать их как образы для классификации. Главными из таких требований являются следующие:

1. инвариантность относительно масштабирования изображения и условий освещения;
2. слабая корреляция различных признаков друг с другом;
3. размерность полученных образов должна быть адекватна размеру используемого словаря ключевых слов.

Приведем четыре типа низкоуровневых признаков, которые представляются наиболее подходящими с точки зрения требуемых свойств, которые были использованы нами для построения системы автоматической аннотации изображений.

1.1 Гистограммы ориентаций градиентов

Гистограммы ориентаций градиентов (Histogram of Oriented Gradient, HoG) позволяют определить наличие локальных объектов на изображении и распознать их форму [10]. Вычисление HoG-дескрипторов производится следующим образом. На изображении вводится прямоугольная сетка, состоящая из малых

пространственных регионов – «клеток», объединенных в большие пересекающиеся области – блоки. Для цветного изображения, представленного тремя цветовыми компонентами в цветовом пространстве RGB, вычисляются матрицы значений градиента с помощью конечных разностей в горизонтальном и в вертикальном направлениях для каждого цветового канала. Затем для каждой точки в качестве модуля и ориентации градиента берутся значения модуля и ориентации в цветовом канале с наибольшим модулем. После этого для каждой клетки блока вычисляется гистограмма направлений градиентов. Гистограммы всех клеток, входящих в один блок, объединяются и образуют вектор коэффициентов блока. Нормализация каждого такого вектора осуществляется путем деления его на нормирующий множитель, равный L_1 норме этого вектора. Затем векторы коэффициентов всех блоков (в порядке от правого верхнего к левому нижнему) объединяются в общий HoG-дескриптор изображения.

1.2 Признак однородности фона

Признак однородности фона может быть полезным, например, для различения таких типов сцен, как пейзаж и портрет, ландшафт и макросъемка [11]. Фоном изображения называют относительно большие по площади связные множества пикселей, составляющих значительную часть всего изображения. Рассмотрим цветное изображение в формате RGB, причем предположим, что элементы его цветовых компонент принимают значения из диапазона $0, 1, \dots, M$. Например, если используется стандартное восьмибитовое представление изображения, то $M = 255$. Для каждой из цветовых компонент изображения определим соответствующие нормализованные гистограммы $h^{col}(i)$, $col = R, G, B$, $i \in 1, \dots, M$. Энтропия каждого цветового канала рассчитывается с помощью выражения:

$$S^{col} = - \sum_{i=0}^M h^{col}(i) \log(h^{col}(i)).$$

На основании полученного значения энтропии определяется признак содержания фона в каждой из цветовых компонент изображения:

$$\pi_{bg}^{col} = 1 - S^{col}/\log M, col = R, G, B.$$

Общий признак содержания фона есть $\pi_{bg} = \pi_{bg}^R \pi_{bg}^G \pi_{bg}^B$. Чем больше значение π_{bg} , тем большую площадь изображения занимает фон. Для получения более информативного представления изображений будем вычислять признак π_{bg} для отдельных прямоугольных областей изображения, полученных разбиением его некоторой сеткой, и объединять полученные значения в вектор признаков.

1.3 Цветовые гистограммы изображений

Цветовые гистограммы позволяют выразить очень важную для зрительного восприятия человека характеристику – распределение цветов на изображении. Большое значение для вычисления цветовых характеристик имеет выбор цветового пространства. Кроме RGB при обработки изображений используются и другие цветовые пространства, например, CIELab. При разработке CIELab предстояла цель создания цветового пространства, изменение цвета в котором

будет линейным с точки зрения человеческого восприятия, т.е. одинаковое изменение значений цвета в разных областях цветового пространства будет производить одинаковое ощущение изменения цвета. Смысл цветовых компонент точки (L, a, b) в пространстве CIELab состоит в следующем. Координата L задает «светлоту» цвета; хроматическая составляющая представлена двумя полярными координатами a и b , показывающими положение цвета в диапазоне от зеленого до пурпурного и от синего до желтого соответственно. Для достижения инвариантности к освещению координата L отбрасывается, а в качестве вектора признаков выбирается двумерная гистограмма из $m_a \times m_b$ ячеек. Число ячеек выбирается эмпирически и не должно быть велико во избежание получения слишком большой размерности признаков. В качестве значений данных параметров можно взять $m_a = m_b = 8$.

1.4 Текстурные признаки изображений

Текстура является очень важным с точки зрения восприятия и распознавания объектов человеком свойством изображений, позволяющим определять свойства поверхностей различных областей изображения. Одним из наиболее информативных текстурных признаков считается введенный в [12] признак, включающий такие характеристики, как грубость, контраст, направленность текстуры. Обозначим через $I_{i,j}$ интенсивность (яркость) пикселя (i, j) анализируемого изображения. Грубость текстуры характеризует размер составляющих ее структурных элементов. Для вычисления этого показателя в каждом пикселе (l, m) изображения рассчитываются шесть значений средней интенсивности пикселей в окне размером $2^k \times 2^k$, $k = 0, 1, \dots, 5$:

$$E_k(l, m) = \frac{1}{2^{2k}} \sum_{i=l-2^{k-1}}^{l+2^{k-1}} I_{i,j}.$$

Затем для каждого пикселя (l, m) вычисляются абсолютные разности между парами средних интенсивностей в непересекающихся окнах в горизонтальном и вертикальном направлениях:

$$\Delta_{k,h}(l, m) = (E_k(l + 2^{k-1}, m) - E_k(l - 2^{k-1}, m)),$$

$$\Delta_{k,v}(l, m) = (E_k(l, m + 2^{k-1}) - E_k(l, m - 2^{k-1})).$$

Для пикселя (l, m) находится то значение $k = k_{best}$, которое соответствует максимальной величине вертикальной разности $\Delta_{k,h}(l, m)$ и горизонтальной разности $\Delta_{k,v}(l, m)$ и в качестве наилучшего размера окна для точки (l, m) берется $k_{best}(l, m) = 2^k$. Тогда значение признака грубости текстуры для изображения размером $N_x \times N_y$ берется равным среднему значению наилучшего размера для всего изображения:

$$\pi_{crs} = \frac{1}{N_x \times N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} k_{best}(i, j).$$

Признак, характеризующий контраст, является показателем того, как уровни серого варьируются в пределах изображения, и в какой степени их распределение смещено к белому или черному. Для вычисления контраста используются центральные моменты второго и четвертого порядков гистограммы уровней серого

изображения (цветные изображения необходимо преобразовать в изображения в серых тонах), и значение признака определяется как:

$$\pi_{con} = \tilde{\sigma}^2 / (\tilde{\mu}_4)^{1/4},$$

где $\tilde{\sigma}$, $\tilde{\mu}_4$ – оценки среднеквадратичного отклонения и эксцесса. Ещё одним признаком текстуры согласно [19] является направленность. Для вычисления направленности строится гистограмма направлений градиента $h_{dir}(\theta)$ и вычисляется степень направленности, связанная с величинами пиков гистограммы:

$$\pi_{dir} = \sum_{i=1}^{n_{peaks}} \sum_{\theta \in \Theta_i} (\theta - \theta_i)^2 h_{dir}(\theta),$$

где n_{peaks} – число пиков гистограммы, больших некоторого порога, θ_i – соответствующее i -тому пику значение угла θ , Θ_i – заданная окрестность i -того пика.

2. Классификация статистических классов

2.1 Понятие статистического класса

Для классификации низкоуровневых признаков изображений с целью получения вторичных признаков будем использовать понятие статистического класса, вводимое в конечномерном случае следующим образом [15–17]. Пусть дано конечное пространство $X = \{x_1, \dots, x_n\}$, а $\mathcal{U} = 2^X$ – алгебра всех его подмножеств. Будем считать, что на алгебре \mathcal{U} задана аддитивная мера V , которую в дальнейшем будем называть объемной мерой. Тогда каждый статистический класс F может быть задан в пространстве X с помощью вероятностной меры P на \mathcal{U} , которая абсолютно непрерывна относительно объемной меры V . Тогда статистический класс может быть задан посредством задания функции плотности вероятности. Поскольку низкоуровневые признаки изображений можно представить с помощью гистограмм, будем считать, что каждый признак – это гистограмма, являющаяся оценкой вероятностного распределения. Напомним, что для конечного случая абсолютная непрерывность означает, что для любого $x_i \in X$ из $P(\{x_i\}) > 0$ следует, что и $V(\{x_i\}) > 0$. Поэтому, можно определить функцию плотности вероятности $h : X \rightarrow \mathbb{R}$, задаваемую формулой:

$$h(x) = \begin{cases} P(\{x\})/V(\{x\}) & \text{if } V(\{x\}) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Тогда вероятность произвольного события $A \in \mathcal{U}$ можно рассчитать по формуле: $P(A) = \sum_{x \in A} h(x)V(\{x\})$. Последнюю сумму можно рассматривать и в качестве интегральной суммы для интеграла Лебега, поэтому последнюю формулу можно записать и так: $P(A) = \int_A h(x)dV$.

В практических приложениях объемная мера может выбираться таким образом, чтобы наилучшим образом разделять интересующие нас статистические классы. В том случае, когда информации для выбора объемной меры недостаточно, обычно полагают, что $V(\{x\}) = c > 0$ для всех $x \in X$, в частности, $c = 1$ или

$c = 1/n$. В последнем случае, очевидно, объемная мера является вероятностной мерой на алгебре \mathcal{U} .

Теоретически **отношение включения** статистических классов вводится с помощью так называемых минимальных событий. В данной статье мы не будем останавливаться подробно на этом вопросе (см. [15–17]). Для практических приложений достаточно знать, как вводится отношение включения с помощью функций принадлежности. Пусть задан статистический класс F , определяемый вероятностной мерой P с плотностью $h_F(x)$. Тогда функции

$$\underline{\mu}_F(x) = \sum_{y \in X | h_F(y) < h_F(x)} P_F(\{y\}) \text{ и } \bar{\mu}_F(x) = \sum_{y \in X | h_F(y) \leq h_F(x)} P_F(\{y\})$$

называются соответственно нижней и верхней функциями принадлежности статистического класса F . Будем считать, что по определению статистический класс F_1 включается в статистический класс F_2 , т. е. $F_1 \subseteq F_2$, если $\underline{\mu}_{F_1}(x) \leq \underline{\mu}_{F_2}(x)$ и $\bar{\mu}_{F_1}(x) \leq \bar{\mu}_{F_2}(x)$ для всех $x \in X$. Далее мы также будем рассматривать теоретико-множественные операции над статистическими классами. Они производятся с помощью операций минимума и максимума:

1. $\underline{\mu}_{F_1 \cap F_2}(x) = \min(\underline{\mu}_{F_1}(x), \underline{\mu}_{F_2}(x))$, $\bar{\mu}_{F_1 \cap F_2}(x) = \min(\bar{\mu}_{F_1}(x), \bar{\mu}_{F_2}(x))$ – это функции принадлежности статистического класса $F_1 \cap F_2$;
2. $\underline{\mu}_{F_1 \cup F_2}(x) = \max(\underline{\mu}_{F_1}(x), \underline{\mu}_{F_2}(x))$, $\bar{\mu}_{F_1 \cup F_2}(x) = \max(\bar{\mu}_{F_1}(x), \bar{\mu}_{F_2}(x))$ – это функции принадлежности статистического класса $F_1 \cup F_2$.

Возможна ситуация, когда статистические классы $F_1 \cap F_2$ или $F_1 \cup F_2$ не порождаются вероятностными мерами. Их смысл будет уточнен далее при рассмотрении задачи классификации статистических классов. Статистические классы можно рассматривать как представление вероятностных распределений с помощью нечетких множеств. В частности, в случае, когда верхняя и нижняя функция принадлежности статистического класса совпадают, то статистический класс называется правильным.

Мера включения $\mu(F_1 \subseteq F_2)$ статистического класса F_1 в статистический класс F_2 вводится, чтобы оценить степень включения класса F_1 в класс F_2 . По определению, значение $\mu(F_1 \subseteq F_2) \in [0, 1]$ и $\mu(F_1 \subseteq F_2) = 1$, если $F_1 \subseteq F_2$. Этот функционал может быть описан разными способами, например, описанными в [15–19]. Далее в этой статье мы будем использовать меру включения, аксиоматически введенную в [18]. Введем вспомогательную функцию:

$$\psi(F_1, F_2) = 0.5 \int_X (\underline{\mu}_{F_2}(x) + \bar{\mu}_{F_2}(x)) dP_1 = 0.5 \sum_{x \in X} (\underline{\mu}_{F_2}(x) + \bar{\mu}_{F_2}(x)) P_1(x).$$

Тогда мера включения может быть определена как:

$$\mu(F_1 \subseteq F_2) = \psi(F_1, F_1 \cap F_2) + 1 - \psi(F_2, F_1 \cup F_2).$$

Отметим, что в последней формуле необходимо рассчитать функции принадлежности статистических классов $F_1 \cap F_2$ и $F_1 \cup F_2$.

2.2 Построение вторичных признаков изображений на основе меры включения

Пусть среди множества статистических классов выделено подмножество классов $\{S_1, S_2, \dots, S_n\}$, которые будем называть эталонными. Тогда классификация статистического класса по эталонным классам заключается в нахождении классифицирующего вектора $(\mu(F \subseteq S_1), \dots, \mu(F \subseteq S_n))$.

Покажем, как можно получить вторичные признаки, используя меры включения статистических классов. Поскольку низкоуровневые признаки изображений можно представить с помощью гистограмм, будем считать, что каждый признак – это гистограмма, являющаяся оценкой вероятностного распределения. Например, признаки, получаемые с помощью направлений градиента, представляют собой множество гистограмм, соответствующих различным положениям сканирующего окна. Признаки грубости изображения и однородности фона можно также представить в виде вероятностного распределения, предполагая, что мы рассчитываем эти признаки для различных положений сканирующего окна и в результате получаем выборку из распределения этого признака по всему изображению. Таким образом, все низкоуровневые признаки извлекаемые из изображений можно представить в виде гистограмм.

Рассмотрим, как строятся эталонные классы. Будем считать, что изображения аннотированы таким образом, что каждое аннотируемое слово относится либо к части изображения, либо к всему изображению в целом. Например, если изображение аннотируется ключевым словом «здание», то для увеличения точности автоматической аннотации следует рассчитывать статистические характеристики не для всего изображения в целом, а только к той его части, где изображено здание. Далее предположим, что мы хотим построить вторичные признаки для зданий. Для этого вначале надо построить эталонные классы, соответствующие ключевому слову здание. Для этого можно использовать наиболее простой принцип обобщения, вычисляя среднюю гистограмму для всех изображений из обучающей выборки, которые аннотируются ключевым словом «здание». Будем считать, что построен эталонный класс S , соответствующий ключевому слову w и заданному низкоуровневому признаку b . Тогда для классификации изображения требуется вычислить гистограммы, соответствующие выбранному низкоуровневому признаку. Пусть данные гистограммы описываются статистическими классами F_1, \dots, F_l . Тогда в качестве вторичного признака будем рассматривать число: $p(w|b) = \max \{\mu(F_1 \subseteq S), \dots, \mu(F_l \subseteq S)\}$. Отметим, что в последней формуле использован максимум, так как логично в качестве вторичного признака выбрать значение меры включения части изображения, которое наиболее всего соответствует ключевому слову w . Подчеркнем, что мера включения имеет вероятностную интерпретацию – это усредненное значение условной вероятности того, что при наблюдении статистического класса F мы будем наблюдать минимальные события эталонного класса S . Таким образом, чем больше значение $p(w|b)$, тем с большей вероятностью изображение аннотируется ключевым словом w .

2.3 Агрегирование вторичных признаков

Будем считать, что на этом этапе у нас есть словарь ключевых слов $W = \{w_1, \dots, w_m\}$ и множество признаков $B = \{b_1, \dots, b_n\}$. Предположим, что после этапа выделения вторичных признаков у нас имеются значения $p(w_i|b_j)$. Дальнейшая обработка вторичных признаков состоит в построении агрегирующих

функций $\varphi_i : [0, 1]^n \rightarrow [0, 1]$, с помощью которых строятся глобальные признаки: $p(w_i) = \varphi_i(p(w_i|b_1), \dots, p(w_i|b_n))$, $i = 1, \dots, n$, которые должны нам давать глобальную оценку релевантности ключевого слова анализируемому изображению по всем признакам.

Напомним необходимые свойства агрегирующих функций [13]. Пусть $\varphi : [0, 1]^n \rightarrow [0, 1]$ – агрегирующая функция, тогда

1. $\varphi(\mathbf{0}) = 0$ и $\varphi(\mathbf{1}) = 1$, где $\mathbf{0} = (0, \dots, 0)$ и $\mathbf{1} = (1, \dots, 1)$;
2. $\varphi(\mathbf{x}) \leq \varphi(\mathbf{y})$ для $\mathbf{x} \leq \mathbf{y}$, где $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, и $\mathbf{x} \leq \mathbf{y}$ если $x_i \leq y_i$ для всех $i \in \{1, \dots, n\}$.

Предположим, что признаки являются независимыми. Тогда в качестве агрегирующей функции можно взять линейную функцию, т.е. $\varphi(\mathbf{x}) = \sum_{i=1}^n a_i x_i$, где

$a_i \geq 0$, $i = 1, \dots, n$, и $\sum_{i=1}^n a_i = 1$. Пусть автоматическая аннотация изображения осуществляется по правилу: изображение аннотируется ключевым словом w_i , если $p(w_i) > \varepsilon_i$, при этом агрегирующие функции φ_i и неотрицательные числа ε_i следует оценивать по обучающей выборке аннотированных изображений следующим образом. Предположим, что у нас есть обучающая выборка, состоящая из N изображений, тогда каждое изображение с номером $k \in \{1, \dots, N\}$ будет описываться вектором вторичных признаков $\mathbf{p}_k = (p_k(w_i|b_1), \dots, p_k(w_i|b_n))$, характеризующих релевантность ключевого слова w_i . Будем кодировать информацию о том, аннотируется или нет изображение с номером k словом w_i числом $\delta_k \in \{-1, 1\}$, полагая, что $\delta_k = 1$ ключевое слово w_i присутствует в аннотации изображения с номером k , и $\delta_k = -1$, в противном случае. Тогда в качестве задачи статистического обучения можно рассматривать проблему нахождения вектора $\mathbf{a} = (a_1, \dots, a_n)^T$ и порогового значения ε таким образом, чтобы число неправильных классификаций, т.е. число истинных неравенств $\delta_k(\mathbf{p}_k \mathbf{a} - \varepsilon) > 0$ для $k = 1, \dots, N$, было минимальным. Такая постановка соответствует классической задаче построения линейного классификатора и может быть решена с помощью одного из известных алгоритмов, например, алгоритма персептрона [14].

Заключение

В данной работе рассмотрена актуальная задача автоматической аннотации изображений. Предложена общая схема решения данной задачи, основанная на использовании низкоуровневых признаков изображений. Рассмотрены ключевые свойства этих признаков и описано построение нескольких типов используемых авторами для решения задачи признаков, обладающих требуемыми свойствами. Далее на основе выделенных низкоуровневых признаков строятся вторичные признаки изображений, которые затем классифицируются для получения окончательных аннотаций. Для построения вторичных признаков в данной работе предлагается использовать подход, основанный на классификации статистических классов. Вводимые в работе статистические классы представляют вероятностное распределение низкоуровневых признаков. Также вводится мера включения статистических классов, на которой основана схема классификации статистических классов, соответствующих аннотируемым изображениям, по эталонным классам, соответствующим ключевым словам. Построение аннотаций производится путем агрегирования вторичных признаков с использованием линейных

решающих функций, построенных в результате обучения алгоритма персептрона. Предложенная схема автоматической аннотации изображений была реализована программно, ниже приведены примеры работы системы:



water
sun
landscape
outdoor
lake



night
tree
landscape
city
sky
light

Список литературы

- [1] Tsai C., Hung C. Automatically annotating images with keywords: a review of image annotation systems. *Recent Patents on Computer Science*. 1 (2008) 55–68.
- [2] Hanbury A. A survey of methods for image annotation. *Journal of Visual Languages & Computing*. 19(5) (2008) 617–627.
- [3] Mori Y., Takahashi H., Oka R. Image-to-word transformation based on dividing and vector quantizing images with words. *Neural Networks*. – 1999. - Vol. 1. - P. 405–409.
- [4] Yavlinsky A. Image indexing and retrieval using automated annotation: Ph.D. thesis. University of London, Imperial College of Science, Technology and Medicine; Department of Computing. – 2007.
- [5] Li J., Wang J. Z. Real-time computerized annotation of pictures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. - 2008. - Vol. 30(6). - P. 985–1002.

- [6] Wang Y., Mei Y., Gong S., Hua X-S. Combining global, regional and contextual features for automatic image annotation. *Pattern Recognition*. - 2009. - Vol. 42. - P. 259–266.
- [7] Jeon J., Lavrenko V., Manmatha R. Automatic image annotation and retrieval using cross-media relevance models. *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. - 2003. - Vol. 1. - P. 119–126.
- [8] Lavrenko V., Choquette M., Croft W. B. Cross-lingual relevance models. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. - 2002. - P. 175–182.
- [9] <http://www.alipr.com> – Automatic Photo Tagging and Visual Image Search Service – 2006.
- [10] Dalal N., Triggs B., Schmid C. Human detection using oriented histograms of flow and appearance. *European Conference on Computer Vision*. 1 (2006) 24–31.
- [11] Abramov S.K., Lukin V.V., Ponomarenko N.N. Entropy based background measure calculation for images searching and sorting in the large collections. *Electronics and Computer Systems*. 2(21) (2007) 24–28.
- [12] Tamura H., Mori S., Yamawaki T. Texture features corresponding to visual perception. *IEEE Trans. On Sys. Man, and Cyb.* 8(6) (1978) 460–473.
- [13] Grabisch M., Pap E., Mesić R., Marichal J.-L. Aggregation functions. Cambridge University Press, Cambridge (2009).
- [14] Tsyplkin Y.Z. Adaptation and learning in automatic systems. Academic Press, Inc. Orlando, FL, USA (1971).
- [15] Броневич А.Г., Каркищенко А.Н. Вероятностные и возможностные модели классификации случайных последовательностей. Таганрог: ТРТУ, 1996. - 196 с.
- [16] Броневич А.Г., Каркищенко А.Н. Теоретико-множественный подход к классификации статистических классов. Автоматика и телемеханика, 1994, №2, с. 78–87.
- [17] Броневич А.Г., Каркищенко А.Н. Обобщение понятия статистического класса и меры возможностного включения. Автоматика и телемеханика, 1997, №6 , 84–94.
- [18] Bronevich A.G., Karkishchenko A.N. Application of possibility theory for ranking probability distributions. Proc. of the European Congress on Intelligent Techniques and Soft Computing. (1997) 310–314.
- [19] Bronevich A.G., Karkishchenko A.N. Fuzzy classification of probability distributions. Proc. of the Fourth European Congress on Intelligent Techniques and Soft Computing. 1 (1996) 120–124.