Сёмов А.А. Мера сходства 3D изображений: конструирование решающей процедуры. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XVI Междунар. научно-техн. конф. – Пенза: ПДЗ, 2016. – С. 37-44.

УДК 004.93

МЕРА СХОДСТВА 3D ИЗОБРАЖЕНИЙ: КОНСТРУИРОВАНИЕ РЕШАЮЩЕЙ ПРОЦЕДУРЫ*

А.А Сёмов

3D IMAGES SIMILARITY MEASURE: DECISION PROCEDURE CONSTRUCTION

A.A. Syemov

Аннотация. Описывается конструирование решающей процедуры для отнесения тестового 3D изображения к одному из ранее выбранных классов. Решающая процедура строится на базе теории гипертриплетных признаков, вычисляемых новым геометрическим методом при сканировании и анализе 3D изображения сеткой параллельных плоскостей (гипертрейс-преобразование). Приводится краткое описание аналитической структуры системы гипертриплетных признаков. Анализируются меры сравнения близости 3D изображений, указываются их преимущества и недостатки. Обосновывается выбор функции расстояния, на основе которой строится правило отнесения изображения к тому или иному классу. Приводится пример решающей процедуры.

Ключевые слова: 3D изображение, гипертриплетный признак, функция расстояния, мера сходства дескрипторов признаков (векторов), решающая процедура.

Abstract. In this article, the decision procedure designing to classify a 3D test image to one of the earlier selected classes is described. A decision procedure is built on a theory basis of hypertriplet features computed new geometrical method when 3D images is scanned and analyzed by parallel planes grid (hypertrace transform). A analytical structure brief description of the hypertriplet features system is provided. Comparison measures of 3D images proximity are analyzed, their advantages and drawbacks are indicated. A choice of distance function is substantiated, on which image classification rule is built. The decision procedure example is provided

Keywords: 3D image, hypertriplet feature, distance function, similarity measure of features descriptors (vectors), decision procedure.

Современное процветание компаний, их расширение и конкурентоспособность невозможны без внедрения в сферу деятельности новейших технологий автоматизации процессов управления и контроля, которые невозможны без интеллектуализации компьютерного анализа и распознавания изображений. Применение технологий трехмерного моделирования, обработки и анализа изображений позволяет поднять качество распознавания на целый уровень по сравнению с применением технологий двумерного представления данных в виде изображений и их обработки [1, 2].

С развитием и усложнением технических устройств актуальной становится также задача вычисления различных метрических свойств объектов. Так, в медицинской диагностике сердечно-сосудистых заболеваний большое значение имеет задача визуального наблюдения структуры живого сердца [3]. Точное определение размеров, формы и других параметров отделов сердца играет важную роль, т.к. помогает врачу оценить степень влияния патологии на структуру и работу

* Работа выполнена при поддержке гранта РФФИ (проект №15-07-04484).

сердца. Получить четкое представление о форме внутреннего и внешнего контура желудочков и поставить правильный диагноз крайне сложно без использования средств 3D моделирования и распознавания. Для этого необходимо, чтобы распознающая система умела определять вид исходного пространственного объекта, его класса.

Аналитическая структура признака 3D изображения

Обозначим через $B(\eta,r)=\{x \mid x^T \cdot \eta = r\}$ касательную к сфере с центром в начале координат, проходящую через заданную точку X и на расстоянии r от начала координат и вектором нормалью $\eta=[\cos\omega\cdot\sin\varphi, \sin\omega\cdot\sin\varphi, \cos\varphi]$, где ω – угол между осью $0x$ и проекцией отрезка OX на плоскость $0xy$, φ – угол между осью $0z$ и отрезком OX.

Исходный 3D объект F сканируется сеткой параллельных плоскостей [4]. Результат пересечения сканирующей плоскости $B(\eta(\omega,\varphi),r)$ с трёхмерным изображением F характеризуется числом: $G = HyperT\,(F \cap B(\eta(\omega,\varphi),r))$. Так как сканирование зависит от трех параметров плоскости $B(\eta(\omega,\varphi),r)$, то в результате вычислений формируется гипертрейс матрица 3ТМ, у которой ось 0ω направлена горизонтально, ось 0φ – вертикально, ось $0r$ – вглубь [5].

Далее, используя построковые и постолбцовые обработки матрицы *3TM*, при помощи функционалов *HyperP*, *HyperΩ* и *HyperΘ* вычисляется признак *Res*(*F*) исходной модели [6]:

$$Res(F) = Hyper\boldsymbol{\Theta} \circ Hyper\boldsymbol{\Omega} \circ Hyper\boldsymbol{P} \circ HyperT(F_{sect}).$$

Как видно из представленной выше формулы, сформированная система гипертриплетных признаков имеет однородную структуру, что говорит об её относительной универсальности. Практически любой признак, отражающий свойство исходного 3D изображения, может быть сконструирован на основе данной формулы.

После того, как задано описание 3D изображение в виде дескриптора признаков (вектора значений), задача распознавания ещё не решена. Необходимо задать правило для отличия одних изображений пространственных объектов от других (вычисление класса изображения). Чтобы распознающая система умела определять класс исходного пространственного объекта, необходимо задать правило отнесения его изображения к тому или иному классу – решающую процедуру, основное ядро которой составляет мера, определяющая сходство двух изображений.

Мера сходства изображений. Постановка задачи

Мера сходства изображений – количественная характеристика соответствия образов согласно выбранному критерию по элементам данных изображений. Данная мера характеризуется функцией расстояния исходного изображения до одного из классов или другого изображения.

Пусть исходные данные представлены в виде матрицы сходства $\|d(x,y)\|$, где $d(x,y)$ – некоторая мера близости между каждой парой дескрипторов признаков x и y пространственных объектов X и Y, а $x_1, \dots, x_n$ и $y_1, \dots, y_n$ – наборы признаков для исходных 3D изображений X и Y соответственно. Хорошо известно, что для любого заданного разбиения объектов на группы и любого $\varepsilon > 0$ можно указать метрику, такую, что расстояния между объектами из одной группы будут меньше ε, а между объектами из разных групп – больше 1 / ε. Построение такого правила отнесения и есть эффективная классификация базы 3D изображений.

Приведём описание основных подходов задания данной меры (функции расстояния):

1. Мера Минковского (степенное расстояние):

$$d(x,y)=\left\{\sum_i (x_i-y_i)^r\right\}^{\frac{1}{p}}.$$

Стоит отметить, что при $r = p = 2$ выражение описывает обычное евклидово расстояние между векторами признаков изображений, которое очень часто используется в различных задачах. Евклидова метрика – наиболее естественная функция расстояния, возникающая в геометрии, отражающая интуитивные свойства расстояния между точками x и y в n-мерном пространстве.

Возможны и другие модификации. При $r = p = 1$ метрика Минковского дает "расстояние городских кварталов" (*манхэттенское расстояние*), которое является просто суммой разностей по координатам: $d(x,y)=\sum_i |x_i-y_i|$. В большинстве случаев эта мера расстояния приводит к таким же результатам, что и обычное расстояние Евклида. Однако отметим, что для нее влияние отдельных больших разностей (выбросов) уменьшается, так как они не возводятся в квадрат.

При $r = p \to\infty$ имеем метрику доминирования (супремум-норма или расстояние Чебышева), которая вычисляется по формуле: $d(x,y)=\max_i |x_i-y_i|$. Это расстояние может оказаться полезным, когда желают определить два объекта как "различные" по какой-либо одной лимитирующей координате (каким-либо одним измерением).

Отметим, что мера Минковского удовлетворяет всем аксиомам расстояния. Она удобна для определения расстояния между двумя точками, например, между точкой наблюдаемых параметров и центром (выборочным средним) класса. Применяется в случае, когда необходимо увеличить или уменьшить вес, относящийся к размерности, для которой соответствующие объекты сильно отличаются.

Тем не менее, она имеет недостатки: не учитывает распределение точек в классе. На примере категориальных значений недостаток этой метрики в том, что пользователи, оценившие один фильм одинаково (других общих фильмов нет), могут быть более схожи, чем оценившие сотню фильмов почти одинаково.

2. Коэффициент корреляции Пирсона:

$$d(x,y)=\frac{n\sum_{i=1}^{n}x_i\cdot y_i-\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)}{\sqrt{\left(n\sum_{i=1}^{n}x_i^{\,2}-\left(\sum_{i=1}^{n}x_i\right)^2\right)\left(n\sum_{i=1}^{n}y_i^{\,2}-\left(\sum_{i=1}^{n}y_i\right)^2\right)}}.$$

Коэффициент корреляции Пирсона – это число d ($-1 \le d(x,y) \le 1$), которое показывает тенденцию двух множеств чисел, идущих попарно, быть схожими. Другими словами, если наблюдается зависимость между векторами, то значение корреляции по модулю будет ближе к 1. Если зависимость линейная, то коэффициент корреляции равен 1, если обратная – то будет равно -1, если нет зависимости – то 0.

Коэффициент корреляции Пирсона отражает зависимость между векторами более-менее эффективно при следующих допущениях:

1) Элементы исследуемых массивов отражают динамику некоторого показателя во времени или однородные свойства некоторого признака по выбранным показателям. Другими словами, если требуется узнать одинаковость направлений роста и убывания в значениях массива согласно их местоположению. Изменение элементов массива местами изменяет значение корреляции.

2) Отсутствуют резкие колебания в элементах массивов.

На примере категориальных значений недостаток этой метрики можно выразить в следующем. Логично предположить, что два пользователя, посмотревшие 100 одинаковых фильмов, больше похожи, даже если они выставили разные оценки, чем два пользователя, посмотревшие три одинаковых фильма и поставившие им одинаковые оценки. Но из формулы мы получим противоположный результат.

3. Вероятностная модель с использованием идеи энтропии или вероятностного распределения имеющихся элементов:

$$d(x,y) = min \left|H(\xi_x) - H(\xi_y)\right|,$$

где $H(\xi) = -\sum_{\xi} p(\xi) \cdot ln(p(\xi))$ – энтропия множества X исходов (значений) случайной величины ξ, $0 \le p(\xi) \le 1$ – её распределение вероятностей.

Недостаток данного подхода заключается в обязательном нахождении и оценке плотности распределения той или иной величины, что не всегда оправдано в реальных практических задачах (особенно, когда имеющая выборка нерепрезентативна).

Другие подходы к заданию функции расстояния строятся на основе данных трёх или их комбинаций.

Решающая процедура для определения класса изображения

Стоит отметить, что мы будем определять меру близости изображений не через их трехмерную модель (массив данных), а через их векторы – дескрипторы признаков. Элементы вектора сами по себе не являются однородными: каждый элемент-признак отражает определённое то или иное свойство 3D изображения. Необходимо использовать тот вид критерия для определения меры близости двух векторов, который менее всего подвержен изменению при изменении входных данных.

При этом стоит помнить, что с точки зрения надежности распознавания не все признаки равноправны: использование одних признаков позволяет получить меньшую вероятность ошибки распознавания, использование других – большую. Поэтому желательно оценивать также информативность каждого признака в отдельности. Кроме того, порядок следования признаков в векторе дескрипторе не должен влиять на функцию расстояния. Вычислять плотность распределения значений элементов данного вектора бессмысленно, так как она будет сильно будет зависеть от выбранных признаков. Вычисление всех видов признаков неприемлемо из-за больших вычислительных затрат.

Учитывая всё вышесказанное (включая предыдущий раздел), была выбрана следующая функция расстояния. Обозначим через t тестовое 3D изображение из какого-либо подмножества B_i. Тогда его k-ый признак будет равен Res_k^t. Расстояние между тестовым 3D изображением и i-ым классом A_i с учётом весовых значений определяется следующим образом:

$$d(t,A_i)=\sum_k p(A_i,k)\cdot\frac{\left|Res_k^t-\mu_k^{A_i}\right|}{\sigma_k^{A_i}}.$$

Распознающая система тестовое изображение t относит к классу A_j, если:

$$d(t,A_j)=\min_i d(t,A_i).$$

Решающая процедура для определения класса 3D изображения построена таким образом, что может как учитывать, так и не учитывать весовые коэффициенты для каждого информативного гипертриплетного признака. Применение весов улучшает информативность функции расстояния, что было доказано экспериментальных путём [7].

Таким образом, важное преимущество гипертриплетных признаков заключается в том, что при опоре на большое их количество применяются простые решающие правила для распознавания 3D изображений. При этом при определении принадлежности тестового изображения учитываются весовые коэффициенты для каждого информативного признака в зависимости от класса 3D изображения. Сформированная матрица весов показывает различающую силу признаков с точки зрения уровня потенциальных ошибок в классификации объектов, что повышает точность распознавания.

Заключение

Разработана процедура сокращения размерности признакового пространства, которая позволяет не только отбирать заданное число информативных признаков, но и присваивать каждому из них весовой коэффициент, обозначающий его различающую силу в зависимости от предъявляемого класса 3D изображения. Опора на минимальный набор эффективных признаков значительно сокращает время работы распознающего алгоритма и повышает его эффективность, что было доказано практически результатами экспериментов, позволяет получать набор информативных признаков с указанием их различающей силы для каждого класса 3D изображений, что повышает точность распознавания.

Библиографический список

1. Song S., Xiao J. Sliding shapes for 3D object detection in depth images // Proceedings of the 13th European Conference on Computer Vision (ECCV 2014). Zurich, Switzerland : Springer, 2014. Part VI. Vol. 8694. P. 634-651.

2. Wang C., Huang K.-Q. VFM: visual feedback model for robust object recognition // Journal of Computer Science and Technology. 2015. Vol. 30. Issue 2. P. 325-339.

3. Алпатов А.В., Калинин Р.Е. Способ количественной визуализации формы правого желудочка сердца человека в целях эхокардиографических исследований // Актуальные вопросы клинической морфологии: сборник научных трудов. Рязань: Изд-во РГМУ, 2000. С. 80–84.

4. Fedotov N.G., Ryndina S.V., Syemov A.A. Trace transform of spatial images // 11th International conference on Pattern Recognition and Image Analasis: New Information technologies (PRIA-11-2013). Conference Proceedings (V. I-II). Samara: IPSI RAS, 2013. V. 1. P. 186-189.

5. Сёмов А.А. Основные методы построения гипертрейс-матриц // XXI век: итоги прошлого и проблемы настоящего плюс : научно-методический журнал. 2015. № 03 (25). Т. 1. С. 69-76. (Технические науки. Информационные технологии).

6. Fedotov N.G., Ryndina S.V., Semov A.A. Trace transform of three-dimensional objects: recognition, analysis and database search // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. 2014. Vol. 24. No. 4. Moscow: Pleiades Publishing, Ltd. P. 566-574.

7. Сёмов А.А. Методы анализа и распознавания трёхмерных изображений на основе стохастической геометрии: дис. ... канд. техн. наук: 05.13.17 // Сёмов Алексей Александрович. Пенза, 2015. 140 с.

Сёмов Алексей Александрович
ООО «КОМЭРФ», г. Пенза, Россия
E-mail: mathematik_aleksey@mail.ru

Syemov A.A.
Ltd «COMEARTH»,
Penza, Russia