

ISSN 2500-316X (Online)

<https://doi.org/10.32362/2500-316X-2020-8-3-7-13>



УДК 004.932

Модель репринта объекта на изображении

А.А. Куликов

МИРЭА – Российский технологический университет, Москва 119454, Россия
@Автор для переписки, e-mail: tibult41@gmail.com

Представлена проблематика распознавания (идентификации) лицевых изображений. Показано различие распознавания и идентификации лицевых изображений. Для решения проблемы идентификации разработана модель репринта объекта на изображении. Данная модель решает проблему посредством представления объекта в 3-мерном виде, что позволяет оценить и сформировать необходимые характеристики объекта в полном объеме, тогда как в 2-мерном виде это сделать невозможно. Модель репринта объекта на изображении может использоваться для формирования репринта любых пространственных объектов. Для обучения модели репринта объекта на изображении используется многослойная нейронная сеть, которая обучается последовательно. Для учета допустимых изменений ракурса, различных помех и разных уровней освещенности разработан локальный детектор для модели идентификации лицевых изображений. Бинарное значение, являющееся результатом обработки модели, представляемое как активация, определяет отношение конкретного изображения к соответствующему классу. Локальный детектор является не только основным элементом модели репринта объекта на изображении, но это еще и отдельная математическая конструкция. Он принимает входные данные в качестве двумерных изображений. Разработанная модель репринта объекта на изображении полностью решает проблему идентификации человека по лицевому изображению в целом в условиях помех и независимо от изменения ракурса.

Ключевые слова: нейронная сеть, распознавание изображений, распознавание образов, модель идентификации, идентификация лицевых изображений.

Для цитирования: Куликов А.А. Модель репринта объекта на изображении. *Российский технологический журнал.* 2020;8(3):7-13. <https://doi.org/10.32362/2500-316X-2020-8-3-7-13>

The model is a reprint of an object in the image

Aleksandr A. Kulikov

MIREA – Russian Technological University, Moscow 119454, Russia

@Corresponding author, e-mail: tibult41@gmail.com

The problem of facial image recognition (identification) is presented. The difference between facial image recognition and identification is shown. To solve the identification problem, a model of object reprint in the image was developed. This model solves the problem by representing the object in 3-dimensional form, which makes it possible to evaluate and form the necessary characteristics of the object in full, whereas in 2-dimensional form, it is impossible to do this. The model of an object reprint in an image can be used to create a reprint of any spatial objects. To train a reprint model of an object in an image, a multi-layer neural network is used, which is trained sequentially. A local detector for the facial image identification model has been developed to account for acceptable changes in angle, various noise, and different light levels. The binary value that is the result of model processing, represented as activation, determines the relation of a particular image to the corresponding class. The local detector is not only the main element of the model of the object reprint in the image, but it is also a separate mathematical construction. It accepts input data as two-dimensional images. The developed model of object reprint on the image completely solves the problem of identifying a person from the facial image as a whole in conditions of interference and regardless of changes in the angle.

Keywords: neural network, image recognition, identification model, facial image identification.

For citation: Kulikov A.A. The model is a reprint of an object in the image. *Rossiiskii tekhnologicheskii zhurnal = Russian Technological Journal*. 2020;8(3):7-13 (in Russ.). <https://doi.org/10.32362/2500-316X-2020-8-3-7-13>

Большое количество работ [1–7], связанных с идентификацией лицевых изображений, не решает саму проблему идентификации, а именно, отношение конкретного лица (класса) к исходному изображению. Проблема распознавания тесно связана с проблемой идентификации. Различие заключается в следующем: распознавание лицевого изображения состоит в нахождении на фотографии или в видеоматериале признаков лица, а идентификация – в соотнесении лицевого изображения, конкретно найденного на фотографии или в видеоматериале с помощью распознавания, с лицевым изображением, хранящимся в системе идентификации. Главная сложность метода идентификации по лицевому изображению состоит в том, что система должна идентифицировать человека при различных «помехах», внешних изменениях лица (например, пластика, косметика и др.) и, самое главное, независимо от изменения ракурса. Для решения этой проблемы в статье предложена модель репринта (представления) объекта (МРО) на изображении, которая использует методы восстановления трехмерных характеристик объекта из его двухмерных характеристик. Эта модель может использоваться для формирования репринта любых пространственных объектов. Она представляет собой последовательность уровней репринта E_0, \dots, E_n , состоящих из отдельных локальных детекторов (ЛД) признаков и входных данных. Для обучения такой модели необходимо использование многослойной нейронной сети. Многослойная нейронная сеть должна обучаться последовательно.

В ходе обработки входных данных результат, полученный на предыдущем уровне, передается на вход последующего уровня для получения итогового результата, который в свою очередь представляется не только в виде значения функции активации ЛД, но также и в виде совокупности параметров объекта обрабатываемым ЛД. В совокупности все это называется входными данными для ЛД.

На вход уровня E_0 необходимо подать данные локальных участков покадровых последовательностей изображений, извлекаемых из видеоизображения. После обучения первого уровня видеофрагменты представляются в виде графов (рис. 1).

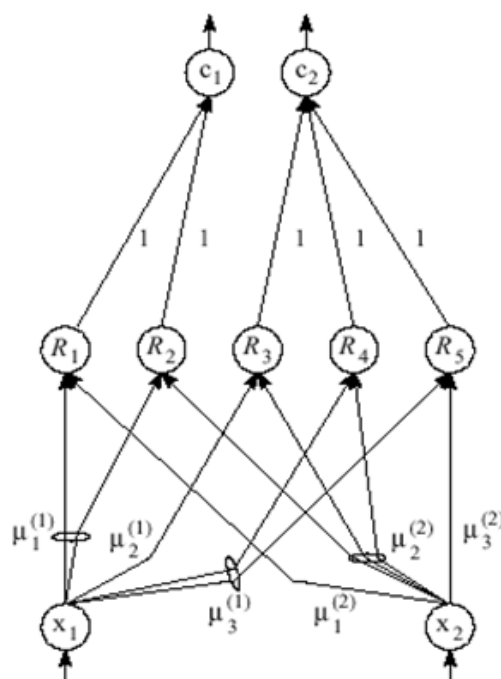


Рис. 1. Представление видеофрагмента в виде графа:

x – изображение на входе, R – набор покадровых последовательностей, c – результат на выходе.

Обученные на первом уровне локальные детекторы, являются узлами графа. Каждый i -ый уровень для $i \neq 0$ осуществляет локальную группировку данных. Все остальные уровни обучаются таким же образом. Обученная модель может представить изображение, извлеченное из видеофрагмента, в виде ограниченного количества ЛД. В каждом локальном детекторе содержится представление комплексного объекта.

Модель репринта объекта на изображении должна содержать:

- количество и локальное расположение детекторов D_i ;
- количество кластеров K_c для каждого уровня E_c ;
- количество уровней репринта E_0, \dots, E_n ;
- внутренние параметры детекторов.

Например, изображения $x_n \in X$ в функции распознавания изображений " $g: X \rightarrow Y$ " представлены в виде вектора длины n атрибута данной функции. Множество классов $y \in Y$, используемых в этой функции, представлены в виде значения этой функции, которое в свою очередь модифицируется под конкретно поставленную задачу.

Для проверки модели распознавания объекта на изображении используем функцию $h: X \rightarrow Y$ для подмножества из множества парных атрибутов и значений $D = \{(x_0, y_0), \dots, (x_m, y_m)\}$,

которая в свою очередь аппроксимирует функцию распознавания g во всей области определения. Входные данные изображения необходимо подать на первый слой модели распознавания, чтобы вычислить h . Далее необходимо последовательное выполнение активации ЛД для каждого следующего уровня модели распознавания лицевых изображений. Бинарное значение, представляемое как активация, является результатом обработки модели и определяет отношение конкретного отдельного изображения к соответствующему классу. Это бинарное значение имеет два состояния:

1 – на выходе модели рассчитывается вероятностная оценка параметров изображения. Для решения задач классификации можно использовать данную модель в исходном виде. Множество классов $y \in Y$ состоит из двух элементов и равно $\{0,1\}$. Функция $g = 1$, если изображение содержит объект, относящийся к необходимому классу.

0 – в противном случае.

Задача идентификации изображения объекта среди множества классов состоит из:

1) отношения последовательной проверки объекта изображения к экземпляру модели идентификации лицевых изображений, т.е. функция $g' = \{g_1, g_j\}$. $g' \in X'$, величины которой – совокупность классов Y' , для i -го класса $Y' = \{y_i, \bigcup_{j \neq i} y_j\}$;

2) для каждого конкретного класса происходит обучение модели.

Схема модели, состоящая из двух уровней репринта, представлена на рис. 2.

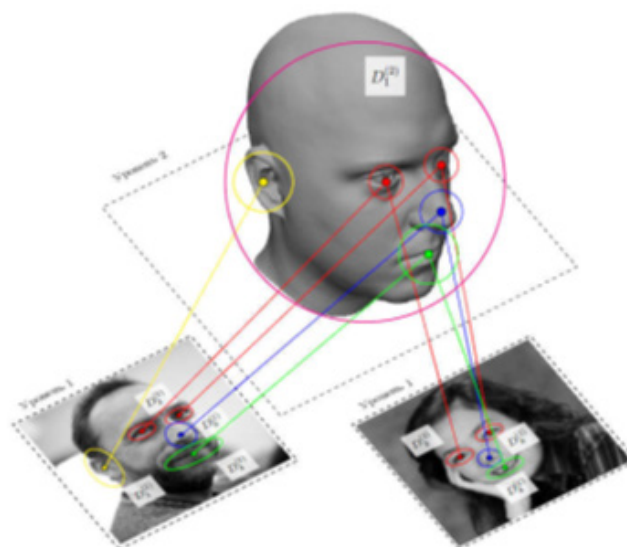


Рис. 2. Модель представления объектов, содержащая два уровня репринта.

Локальный детектор является не только основным элементом модели репринта объекта на изображении, но это еще и отдельная математическая конструкция. Его роль близка к роли нейронов искусственной нейронной сети. Результаты обработки информации в ЛД на предыдущем уровне поступают на вход локального детектора. Выходные данные локального детектора отправляются на следующий уровень модели.

Дадим определение локального детектора. Допустим, существует некоторая совокупность n трехмерных точек с заданными векторными координатами. Такая совокупность является матрицей размером $3 \times n$, которая в свою очередь является множеством B .

Проекцию трехмерных точек на плоскости, можно описать как $Br(B) = CB$, где C является матрицей. Исходя из определения локального детектора, результат решения такой матрицы будет состоять из массива точек b с координатами (x, y) . Такое определение ЛД четко описывает структуру, поведение и результат работы детектора в модели идентификации лицевых изображений. $Mr(M, B) = M$ – операция аффинной трансформации в трехмерном пространстве. M – матрица. Проекция объекта – $b_0 = Br(B)$. Приведем пример. Допустим, существует некоторая совокупность некоторых матриц трансформаций M_1, \dots, M_m и некоторая совокупность проекций b_1, \dots, b_m , таких, что $b_j = Br(Tr(M_j, B))$. Тогда модель идентификации лицевых изображений для таких совокупностей будет использовать функции $D_I(b_j) = b_0$; $DM(b_j) = M_j$, что и является в данном случае локальным детектором. ЛД использует совокупность представлений объекта в трехмерном виде. Для любого представления объекта он способен не только определить представление измененного объекта, но и определить изначальную форму объекта.

Универсальность определения ЛД позволит избавиться от формы объекта, когда использование формы объекта не нужно. Например, когда $b_i = T_i b$, $b_j = T_j B$, b_i, b_j – представление объекта, то верна следующая формула:

$$\begin{cases} D_I(b_i) = D_I(b_j) \\ \frac{D_T(b_i)}{D_I(b_j)} = \frac{T_i}{T_j} \end{cases} \quad (1)$$

Дополним определение ЛД для объектов, которые относятся к разным классам.

1. Пусть B_0, \dots, B_n – совокупность объектов.
2. Y_0, \dots, Y_n – совокупность классов.
3. $y(B) = Y$ – определение принадлежности трехмерного объекта к конкретному классу.

Такое определение локального детектора в модели идентификации лицевых изображений позволяет точно сопоставить трехмерный (3D) объект определенному классу.

Приведем функцию однокомпонентной классификации y'_k для некоторого класса Y_k такую, что

$$y'_k(B_i) = \begin{cases} 1, y(b_j) = Y_k, \\ 0, \text{в противном случае.} \end{cases} \quad (2)$$

Тогда $(D_I^k D_T^k)$ – дискриминативный ЛД для класса Y_k , если он задан функциями D_I^k, D_T^k , такими, что

$$\begin{cases} D_I^k(B_i) = y'_k(B_i), \\ D_T^k(B_i) = \begin{cases} T_i, D_I^k(B_i) = 1, \\ 0, \text{в противном случае.} \end{cases} \end{cases} \quad (3)$$

Локальный детектор должен выполнять:

- 1) распознавание объекта,
- 2) определение конкретной трансформации применимой к изображению.

Для оценки эффективности работы ЛД дадим определение понятию функции ошибки восстановления трансформации:

$$J(D_T, b, T) = \frac{1}{2} \sum_{i=1}^m |D_T(b^{(i)}) - T^{(i)}|^2. \quad (4)$$

ЛД принимает входные данные в виде двумерных (2D) изображений. Двумерные изображения, а, следовательно, и сам локальный детектор, не имеет данных о векторных свойствах объекта, а именно, о дополнительной координате Z . В случае, если бы координата Z была известна, то определение трансформации объекта была бы простой задачей, не требующей специализированных моделей с использованием локальных детекторов.

Для данной особенности ЛД существуют следующие признаки:

– Возьмем необходимое количество трансформированных объектов лицевых изображений b_i, b_j , которым отвечают T_i, T_j . Для случая, когда $b_i \approx b_j$, но $T_i \neq T_j$ ЛД не сможет корректно восстановить измененный объект по изображению. Таким образом, $D_T(b_i) \approx D_T(b_j)$.

– Когда 3D объекты B и Q , относящиеся к разным классам Y_1, Y_2 , дают похожие проекции $b \approx q$, такой же результат получается для функции идентификации ЛД, т.е. $D_I(b) \approx D_I(q)$.

Есть такие вариации трансформации (T) и изображений объекта, для которых нельзя построить ЛД ($J \rightarrow 0$). Негативно сказываются любые возможные отклонения 3D форм объектов лицевых изображений на различные проекции (лицевых изображений), момент изменчивости лицевого изображения объекта увеличивает среднее значение ошибки в пределах одного класса. Возможно, существует отрицательная корреляция между точностью восстановления измененного объекта изображения, размером лицевого изображения и множества лицевых изображений. Репринт полного объекта лицевого изображения с помощью только одного ЛД является сложной задачей, которая превосходит по сложности задачу распознавания.

На первом уровне модели представлен ЛД $D_I(0)$, ансамбли и значения которого различны для каждого конкретного объекта лицевого изображения, который может реагировать на участки объекта лицевого изображения при различном ракурсе. Данные из первого уровня модели локального детектора необходимы для обучения высокоуровневого ЛД $D_j^{(1)}$, находящим представлением для композиций ЛД, а именно, 3D карта признаков – ЛД первого уровня, на которой локальные детекторы размещены в соответствии с данными эквивариантных трансформаций. Принадлежность объекта к категории определяет функция идентификации ЛД второго уровня.

Таким образом, предложенная модель репринта объекта на изображении полностью решает проблему идентификации человека по лицовому изображению в целом в условиях помех и независимо от изменения ракурса.

Литература:

1. Коротков А. Е., Трифонова Е. Е. Алгоритм расчета расстояния Левенштейна с пороговым значением. *Естественные и технические науки*. 2012;1:317-322.
2. Куликов А.А., Демкин Д.В., Мелков А.Е. Анализ влияния максимальной степени сжатия изображения лица на результат распознавания лица. *Перспективы науки*. 2014;3(54):104-108.
3. Камышинский В.И., Смирнов Д.А. Нейронные сети и их применение в системах управления и связи. М.: Горячая линия - Телеком, 2002. 94 с. ISBN 5-93517-094-9
4. Хайкин С. Нейронные сети: полный курс: пер с англ. М.: Вильямс, 2008. 1104 с. ISBN 978-5-8459-0890-2

5. Брилюк Д.В., Старовойтов В.В. Распознавание человека по изображению лица и нейросетевые методы: обзорный препринт. Минск: Ин-т техн. кибернетики НАН Беларуси, 2001. 54 с.
6. Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах: учебн. пособие для ВУЗов. Киев: Издательский дом "Слово", 2008. 344 с.
7. Старовойтов В.В. Локальные геометрические методы цифровой обработки и анализа изображений. Минск: Ин-т техн. кибернетики НАН Беларуси, 1997. 284 с.
8. Куликов А.А. Алгоритм нейронной сети NEFCClass M и реализация алгоритма в среде Matlab. *Информационные технологии моделирования и управления*. 2013;3(81):262-268.
9. Куликов А.А. Система автоматической идентификации изображения лица персоны по видеоизображению. Материалы межвуз. студен. науч.-практ. конф. Карьера и образование – 2013. М.: Изд-во МГОУ имени В.С. Черномырдина, 2013. С. 99-100.
10. Балдин А.В., Елисеев Д.В. Адаптируемая модель данных на основе многомерного пространства. *Наука и образование*: [электронный журнал]. 2010. № 10. <http://technomag.edu.ru/doc/161410.html>
11. Елисеев Д.В., Балдин А.В. Алгебра многомерных матриц для обработки адаптируемой модели данных. *Наука и образование*: [электронный журнал]. 2011. № 7. <http://technomag.edu.ru/doc/199561.html>
12. Korotkov A. Database index for approximate string matching. In: Proc. 4th Spring/Summer Young Researchers' Colloquium on Software Engineering. SYRCoSE '10. 2010. P. 136-140. <https://doi.org/10.15514/syrco-se-2010-4-27>
13. Etemad K., Chellapa R. Discriminant Analysis for Recognition of Human Face Images. *J. Opt. Soc. Am. A*. 1997;14(8):1724-1733. <https://doi.org/10.1364/JOSAA.14.001724>

References:

1. Korotkov A.E., Trifonova E.E. Algorithm for calculation of the Levenshtein distance with a threshold value. *Estestvennye i tekhnicheskie nauki = Natur. Tech. Sci.* 2012;1:317-322 (in Russ.).
2. Kulikov A.A., Demkin D.V., Melkov A.E. The Analysis of the Maximum Face Compression Impact on Face Recognition Result. *Perspektivy nauki = SCIENCE PROSPECTS*. 2014;3(54):104-108 (in Russ.).
3. Kamyshinskii V.I., Smirnov D.A. *Neironnye seti i ikh primeneniye v sistemakh upravleniya i svyazi* (Neural networks and their application in control and communication systems). Moscow: Goryachaya liniya – Telekom; 2002. 94 p. (in Russ.). ISBN 5-93517-094-9
4. Khaikin S. *Neironnye seti: polnyi kurs* (Neural networks: full course). Moscow: Vil'yams; 2008. 1104 p. (in Russ.). ISBN 978-5-8459-0890-2
5. [Haykin S. Neural Networks: A Comprehensive Foundation. NJ: Prentice Hall; 1998. 798 p.]
5. Brilyuk D.V., Starovoitov V.V. *Raspoznavanie cheloveka po izobrazheniyu litsa i neirosetevye metody: obzorny preprint* (Facial recognition and neural network methods: a review preprint). Minsk: Institute of Tech. cybernetics of NAS of Belarus; 2001. 54 p. (in Russ.).
6. Zaichenko Yu.P. *Nechetkie modeli i metody v intellektual'nykh sistemakh: uchebn. posobie dlya VUZov* (Fuzzy models and methods in intelligent systems). Kiev: Slovo Publishing House; 2008. 344 p. (in Russ.).
7. Starovoitov V.V. *Lokal'nye geometricheskie metody tsifrovoi obrabotki i analiza izobrazhenii* (Local geometric methods of digital image processing and analysis). Minsk: Institute of Tech. cybernetics of NAS of Belarus; 1997. 284 p. (in Russ.).
8. Kulikov A.A. NEFCclass M neural network algorithm and implementation of the algorithm in Matlab. *Informatsionnye tekhnologii modelirovaniya i upravleniya = Information technologies modeling and management*. 2013;3(81):262-268 (in Russ.).
9. Kulikov A.A. A system for automatic identification of a person's face image from a video image. In: Proc. Career and education-2013. Moscow: MGOU Publishing House; 2013. P. 99-100 (in Russ.).
10. Baldin A.V., Eliseev D.V. Adaptive data model based on the multidimensional Electron space. *journal. Nauka i obrazovanie = Science & education* [electronic journal]. 2010. № 10 (in Russ.). <http://technomag.edu.ru/doc/161410.html>
11. Eliseev D.V., Baldin A.V. Algebra of multidimensional matrices for processing an adapted data model. *Nauka i obrazovanie = Science & education* [electronic journal]. 2011. № 7 (in Russ.). <http://technomag.edu.ru/doc/199561.html>
12. Korotkov A. Database index for approximate string matching. In: Proc. 4th Spring/Summer Young Researchers' Colloquium on Software Engineering. SYRCoSE '10. 2010. P. 136-140. <https://doi.org/10.15514/syrco-se-2010-4-27>
13. Etemad K., Chellapa R. Discriminant Analysis for Recognition of Human Face Images. *J. Opt. Soc. Am. A*. 1997;14(8):1724-1733. <https://doi.org/10.1364/JOSAA.14.001724>

Об авторе:

Куликов Александр Анатольевич, кандидат технических наук, старший преподаватель кафедры инструментального и прикладного программного обеспечения Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). ORCID - 0000-0002-8443-3684

About the author:

Alexander A. Kulikov, Cand. Sci. (Engineering), senior lecturer, Department of Instrumental and Applied Software, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow 119454, Russia). ORCID - 0000-0002-8443-3684