

Метод активного скелета в задаче распознавания формы изображений

Ирина Бакина, Леонид Местецкий, Арчил Цискаридзе *
Московский государственный университет имени М.В. Ломоносова, Москва, Россия
Московский физико-технический институт, Москва, Россия
irina_msu@mail.ru, l.mest@ru.net, achikotsi@gmail.com

Аннотация

В работе рассматривается аналог метода активных контуров для задачи сравнения формы изображений. Считается, что тестовые объекты представляют собой замкнутые односвязные области; для описания эталонов используется модель гибкого объекта.

Ключевые слова: Метод активных контуров, активный скелет, гибкий объект, распознавание формы.

1. ВВЕДЕНИЕ

В работе предлагается подход к распознаванию формы изображений, основанный на идеях метода активных контуров[1].

Описание эталонного образца в виде гибкого объекта является удобной формой его представления [2], которая включает в себя описание циркулярного графа (или циркуляра) объекта с заданием множества допустимых деформаций. В двумерном случае циркулярный граф представляет собой осевой граф (или скелет), с каждой вершиной которого связан круг, отражающий ширину объекта в точке — центре круга. В этой модели силуэт объекта рассматривается как замкнутая область, ограниченная огибающей семейства кругов циркуляра. С помощью применения трансформаций к элементам скелета происходит моделирование различных движений объекта и его частей, которые влекут за собой изменение силуэта объекта. Таким образом, можно пытаться "шевелить" скелет эталона с целью получения наилучшего совмещения силуэта гибкого объекта с тестовым изображением. Указанная идея переносится также на трехмерный случай, где циркулярный граф является пространственным. Здесь уже появляется задача распознавания формы объекта по его двумерному изображению.

Мотивация для выбора термина "активный скелет" объясняется двумя моментами, определяющими предлагаемый подход и аналогию с методом активных контуров. Они сводятся к следующему:

- Осуществляется трансформация эталона под измеренный тестовый объект, а не наоборот.
- Применение деформаций производится с целью подгонки форм эталонного и тестового образцов. При этом в случае активного контура форма описывается своей границей, а в случае активного скелета — осевым графом циркуляра.

*Работа выполнена при финансовой поддержке РФФИ, проекты № 08-01-00670, № 08-07-00305 и № 08-07-00270.

2. ПОСТАНОВКА ЗАДАЧИ

Имеется некоторая база эталонных объектов, которая формируется offline, может проходить некоторую предварительную обработку экспертами и представляет собой совокупность моделей реальных объектов. Например, в случае распознавания жестов в качестве такой базы может быть рассмотрена азбука жестов глухонемых.

Распознаваемый на изображении тестовый образец представляет собой замкнутую односвязную область, для которой необходимо подобрать наиболее похожий эталон из базы.

Обозначим исходный набор эталонных образцов, заданных в форме гибкого объекта, через $E = \{E_i\}_{i=1}^n$, где n — общее число эталонов. Пусть $V_i = \{v_i^\alpha\}_{\alpha \in A}$ — совокупность возможных трансформаций i -ого эталона, а $v_i^\alpha(E_i)$ — гибкий объект, получающийся в результате применения к эталону E_i трансформации $v_i^\alpha \in V_i$ (здесь $A \subset \mathbb{R}$ — множество индексов). Пусть T_i — осевой граф циркуляра эталонного образца, а S_i — его силуэт. Тогда каждый эталон может быть описан следующей парой:

$$E_i = (T_i, V_i) \quad (1)$$

Пусть F — распознаваемая фигура. Введем функцию сходства $\mu(E_i, F)$ исходного эталонного и предъявляемого тестового образцов (аналог функции энергии в методе активных контуров). Тогда в рассмотренных обозначениях задача минимизации имеет вид:

$$(i^*, v_{i^*}^{\alpha^*}) = \underset{v_i^\alpha \in V_i}{\operatorname{argmin}} \mu(v_i^\alpha(E_i), F) \quad (2)$$

Здесь i^* — номер ближайшего эталона, а $v_{i^*}^{\alpha^*}$ — трансформация, при которой достигается его наилучшее совпадение с распознаваемой формой.

Будем считать, что множество трансформаций $V = \{V_i\}_{i=1}^n$ для всех эталонных объектов определяется совокупностью параметров $\vec{a} = (a_1, \dots, a_k)$, т.е. $V = \{v(\vec{a})\}_{\vec{a} \in D}$, где $D = \bigcup_{i=1}^n D_i \subset \mathbb{R}^k$, а $V^i = \{v(\vec{a})\}_{\vec{a} \in D_i}$.

При этом каждый тестовый объект описывается набором соответствующих параметров $\vec{a}_i = (a_1^i, \dots, a_k^i)$: $E_i = E_i(\vec{a}_i)$. Тогда задача минимизации (2) сводится к нахождению похожего эталона E_{i^*} и параметров его трансформации \vec{a}_{i^*} таких, что:

$$(i^*, \vec{a}_{i^*}) : \mu(v(\vec{a}_{i^*}), F) = \min_{i=1, \dots, n} \min_{\vec{a}^i \in D_i} \mu(v(\vec{a}^i), F) \quad (3)$$

3. ДЕКОМПОЗИЦИЯ ЗАДАЧИ

Переход от задачи в форме (2) к ее представлению в виде (3) позволяет рассматривать процесс глобальной оптимизации как совокупность двух подзадач.

Задача нахождения внешнего минимума является задачей классификации тестового объекта. Вначале определяется эталонный образец E_i и соответствующее его набору деформаций множество D_i . В рамках этого множества осуществляется нахождение внутреннего минимума.

Задача нахождения внутреннего минимума — задача подгонки. При ее решении происходит поиск множества трансформаций активного скелета с целью получения наилучшего совпадения с тестовым образцом. Так как каждая трансформация описывается набором параметров $\vec{a} = (a_1, \dots, a_k)$ и $v = v(\vec{a})$, а каждый эталон задается конкретным набором их значений $E_i = E_i(\vec{a}_i)$, то задача подгонки сводится к задаче поиска локального минимума функции $\mu(v(\vec{a}), F)$ при начальном приближении \vec{a}_i .

4. МЕТОД АКТИВНОГО СКЕЛЕТА

Как отмечалось выше, задача нахождения внутреннего минимума в (3) является задачей подгонки, которую предлагается решать методом активного скелета.

Рассмотрим двухмерный случай, когда имеется двухмерный гибкий объект (эталон) и сравниваемая фигура. Если совместить силуэт эталона с фигурой (рисунок 1а), то они сильно отличаются, хотя являются двумя представлениями ладони одного человека. Если же незначительно "пошевелить" скелет эталона,

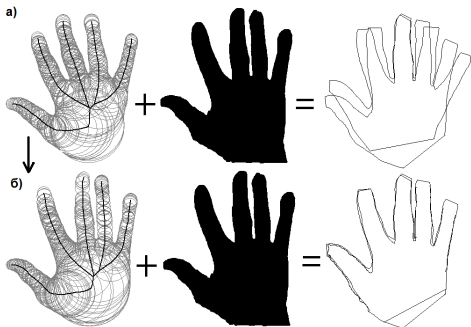


Figure 1: Гибкий объект (левая колонка), сравниваемая фигура (в центре) и результат их совмещения (правая колонка).

моделируя движение пальцев руки, то силуэт эталона и сравниваемое изображение практически совпадут (рисунок 1б).

В трехмерном случае наблюдается похожая ситуация. Предположим, что имеется пространственная эталонная поза человека, описываемая моделью пространственного гибкого объекта. Пусть человека попросили повторить эту позу, после чего сделали ее снимок. В результате, если рассмотреть силуэт гибкого объекта в плоскости камеры и на снимке, то мы увидим их значительное отличие (рисунок 2а). Однако, как и в двухмерном случае, путем "шевеления" скелета пространственного гибкого объекта-позы, можно

добиться хорошего совпадения его силуэта со снимком (рисунок 2б).

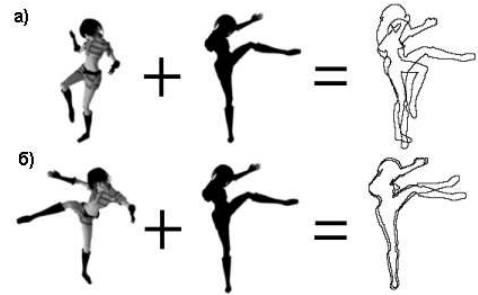


Figure 2: Пространственная поза (левая колонка), сравниваемая фигура (в центре) и результат совмещения их силуэтов (правая колонка).

В данной работе рассматриваются примеры приложения метода активного скелета к двум задачам сравнения формы. Это двухмерная задача распознавание формы ладони человека и трехмерная задача определения позы человека.

5. РАСПОЗНАВАНИЕ ФОРМЫ ЛАДОНИ

Имеется база эталонных бинарных изображений ладоней группы людей. Считается, что эталоны являются "хорошими" изображениями — на них отсутствуют артефакты (кольца и браслеты, длинные ногти, "склеенные" пальцы и т.д.). В то же время, распознаваемые изображения могут обладать ими.

База эталонных ладоней формируется в offline-режиме и содержит в себе модели эталонных изображений ладоней в форме гибкого объекта. Таким образом, имеем $E = \{E_i\}_{i=1}^n$, где n — число эталонных моделей. Тестовые изображения ладоней являются бинаризованными изображениями, полученными с web-камеры.

5.1 Ладонь как гибкий объект

На рисунке 1 в левом столбце представлены циркулярные ладоней. Чтобы задать ладонь в форме гибкого объекта, необходимо дополнительно указать группу деформаций V_i . Определим ее как множество допустимых поворотов пальцев вокруг их оснований (за исключением большого пальца).

Приведение ладони в стандартное положение и процедура нахождения оснований пальцев описаны в работе [3]. На рисунке 3 найденные основания пальцев обозначены звездочками.

Определим параметры трансформации. Рассмотрим углы между осями пальцев и осью среднего пальца руки. На рисунке 3 это углы $\alpha_1, \alpha_2, \alpha_3$ и α_4 . Тогда $E_i = E_i(\vec{\alpha})$, $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. Возможные значения углов поворота пальцев определяют множество $D_i \subset R^4$, а множество трансформаций есть $V_i = V_i(\vec{\alpha})$, где $\vec{\alpha} \in D_i$.

5.2 Сравнение ладоней

Для тестового изображения ладони, также как и для эталонного, можно построить циркулярный граф и привести его в стандартное положение.

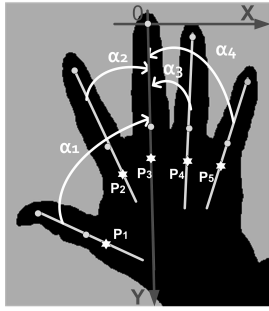


Figure 3: Определение параметров ладони.

По циркулярам эталонного и тестового изображений ладоней далее строится силуэт каждого из них. Через точки поворота указательного пальца и мизинца на эталоне проводится прямая линия, и части силуэтов, лежащие ниже этой прямой, отсекаются.

В качестве функции энергии активного скелета рассматривается площадь симметрической разности:

$$\mu(E_i, F) = \text{Area}(S_i \setminus F) + \text{Area}(F \setminus S_i) \quad (4)$$

При применении трансформаций энергия скелета меняется, и задача подгонки для (3) имеет вид:

$$\vec{\alpha}_i^* : \mu(v_{\vec{\alpha}_i^*}(E_i), F) = \min_{\vec{\alpha}_i \in D_i} \mu(v_{\vec{\alpha}_i}(E_i), F) \quad (5)$$

При сравнении ладоней осуществляются только попарные сравнения пальцев: мизинца с мизинцем, безымянного пальца с безымянным и т.д. То есть этап поиска начального приближения присутствует неявно.

Минимизация (5) выполняется за один шаг — мы измеряем углы между пальцами на тестовом образце, а затем "шевелим" пальцы эталона с целью получения не нем тех же самых углов.

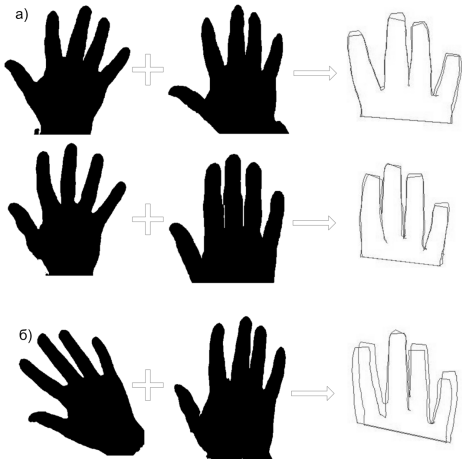


Figure 4: Сравнение ладоней: а) одного человека и б) разных людей.

На рисунке 4 представлен пример сравнения эталонных ладоней (первая колонка) с тестовыми (вторая колонка). В правой колонке представлен результат наилучшего совмещения их силуэтов.

6. РАСПОЗНАВАНИЕ ПОЗЫ ЧЕЛОВЕКА

Фигуру человека можно приближенно рассматривать как совокупность шарнирно закрепленных цилиндрических тел в пространстве, а движение моделировать путем вращениями этих тел вокруг точек крепления. Поза человека — взаимное расположение твердых частей, составляющих его фигуру. Будем считать, что пространственные формы этих тел уже известны и представлены моделью пространственного гибкого объекта.

Предполагается, что имеется пара откалиброванных относительно друг друга web-камер, и эти камеры позволяют с достоверностью выявлять лишь силуэты представленных на изображении объектов. Требуется подобрать такое "шевеление" исходного трехмерного объекта, чтобы его силуэт был похож на каждое из изображений, полученных с камер.

6.1 Метод активного скелета в пространстве

Будем описывать фигуру человека как пространственный гибкий объект $G = (\bigcup_{i=1}^K C_k, V)$ состоящий из $C_k = (O_{k,\alpha}, r_{k,\alpha}), \alpha \in A_k : k = 1, \dots, K$ твердых частей, каждая из которых является семейством сфер с центрами $O_{k,\alpha}$ и радиусами $r_{k,\alpha}$.

Шевеление пространственного гибкого объекта представим в виде совокупности поворотов R_k и переносов t_k , соответствующих его частям C_k . Шарнирные крепления можно записать в виде $(a_l, i_l, b_l, j_l) : \{a_l, b_l\} \in R^3, \{i_l, j_l\} \in 1, \dots, K, l = 1, \dots, L$. В результате имеем L уравнений: $R_{i_l} a_l + t_{i_l} = R_{j_l} b_l + t_{j_l}$.

Задачу распознавания позы по двум силуэтам (F_1, F_2) запишем в виде:

$$(R_i^*, t_i^*) = \underset{\substack{R_i; t_i \in R^3 \\ R_{i_l} a_l + t_{i_l} = R_{j_l} b_l + t_{j_l} \\ i=1, \dots, K, l=1, \dots, L}}{\operatorname{argmin}} \mu\left(\bigcup_{i=1}^K (R_i C_i + t_i), (F_1, F_2)\right) \quad (6)$$

Здесь μ сумма Хаусдорфовых мер различия между силуэтами.

Основная трудность заключается в том, что на плоских силуэтах могут присутствовать окклюзии, и задача (6) имеет много локальных минимумов. Данная трудность может быть преодолена путем предварительной классификации наблюдаемых силуэтов — в случае окклюзий создается база эталонов для их различных конфигураций. В данной задаче эталоном является пространственная поза $v(G), v \in V$.

Для тестовой пары силуэтов, задача минимизации (6) решается в два этапа:

- Определение наиболее похожей конфигурации окклюзии в базе эталонов.
- Осуществление подгонки выбранного на предыдущем шаге эталона.

Для сравнения пары скелетов из рисунка 5а используем метод выделения изоморфной части [4] (похожий подход предложен в [5]), после чего пара скелетов имеет изоморфные оси (рисунок 5б).

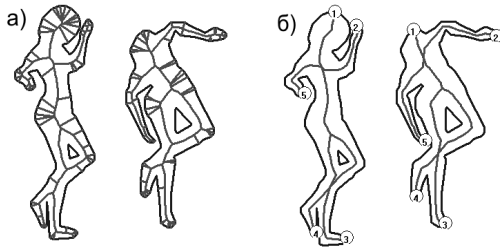


Figure 5: а) Скелеты силуэтов для разных поз и б) общая изоморфная часть скелетов.

Каждая точка силуэта может быть задана в системе координат скелета таким образом, что при деформации скелета ее координаты не меняются. Зная изоморфизм скелетов, можно каждой точке одного силуэта сопоставить точку на другом силуэте, требуя, чтобы у них были одинаковые координаты. Это позволяет строить начальное приближение.

Весь цикл выглядит следующим образом:

- Проецируем пространственный эталон на плоскости камер.
- Строим скелеты для полученных силуэтов.
- Строим изоморфизм между скелетом эталона и скелетами тестовых силуэтов.
- По изоморфизму скелетов строим отображение точек силуэтов.
- Определяем переход проекций осей пространственного эталона при преобразовании силуэтов. По двум проекциям с помощью эпиполярной геометрии [6] строим новые оси в пространстве, получая пространственную деформацию осей эталона.
- Полученную деформацию берем как начальное приближение и запускаем методы локальной минимизации.

6.2 Схема точной подгонки

Поскольку точки каждой твердой части должны сохранять расстояния между собой, скорректируем начальное приближение, поставив задачу определения поворота R_i и переноса t_i для каждой i -ой части $i = 1..K$ как:

$$(R_i^*, t_i^*) = \underset{R_i}{\operatorname{argmin}} \sum_{\alpha \in A_i} |R_i O_{i,\alpha} + t_i - O'_{i,\alpha}|^2 \quad (7)$$

Здесь $O'_{i,\alpha}$ — оси начального приближения.

Наличие шарнирных точек накладывает на R_i^*, t_i^* дополнительные ограничения: $R_i a_i + t_i = R_j b_i + t_j$. Задачу точной подгонки запишем как:

$$(R_i^*, t_i^*) = \lim_{\lambda \rightarrow \infty} \underset{R_i}{\operatorname{argmin}} \mu \left(\bigcup_{i=1}^K (R_i + t_i), F \right) + \lambda \sum_{l=1}^L |R_{i_l} a_l + t_{i_l} - R_{j_l} b_l - t_{j_l}|^2 \quad (8)$$

Таким образом, получив скорректированное начальное приближение, можно запустить градиентный метод для задачи (8).

7. ВЫВОДЫ

В работе рассмотрена общая постановка задачи сравнения формы объектов в случае, когда эталоны представлены моделью гибкого объекта, а тестовые образцы являются замкнутыми односвязными областями. Проведена декомпозиция задачи на две подзадачи — классификации и минимизации (задача классификации актуальна только для случая присутствия окклюзий). Рассмотрен подход к решению обеих подзадач. Для решения задачи минимизации предложен метод активного скелета как аналог метода активных контуров. Основная идея подхода состоит в "шевелении" скелета эталона с целью получения наилучшего совпадения его силуэта с силуэтом тестового экземпляра. Приведены примеры приложения этого метода к задачам сравнения формы для случаев, когда эталонный объект задан в двухмерном и трехмерном пространствах.

8. REFERENCES

- [1] M. Isard A. Blake, "Active contours: The application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion," Springer, 2000.
- [2] Л. Местецкий, "Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры," Москва: ФИЗМАТЛИТ, 2009.
- [3] Л. Местецкий И. Бакина, "Метод сравнения формы ладоней при наличии артефактов," Математические методы распознавания образов (ММРО-14), 2009.
- [4] A. Okhlopkov L. Domakhina, "Shape comparison based on skeleton isomorphism," VISAPP, pp. 237–242, 2009.
- [5] A. L. Yuille S. C. Zhu, "Forms: A flexible object recognition and modeling system," IJCV, vol. 20, no. 3, pp. 187–212, 1996.
- [6] A. Zisserman R. Hartley, "Multiple view geometry in computer vision," Cambridge University Press, 2004.

ACTIVE SKELETON FOR IMAGE FORM RECOGNITION

Abstract

Analogy to the method of active contours is presented for the problem of form recognition in the case when etalon objects are described in terms of flexible object, while test objects are closed 1-connected domains.

Keywords

The method of active contours, active skeleton, flexible object, form comparison.