

Построение дополненной реальности на примере создания виртуальной примерочной.

77-48211/482783

08, август 2012

Девайкин П. А., Шикуть А. В.

УДК.004.021

Россия, МГТУ им. Н.Э. Баумана

p.devaikin@gmail.com

Введение

В настоящее время технологии компьютерного зрения применяются для решения широкого круга задач. Компьютерное зрение основано на алгоритмах распознавания образов. Задача распознавания стала популярной еще в 1960-х годах. Однако возможность создания полноценных систем компьютерного зрения, решающих прикладные задачи, появилась только недавно одновременно с распространением персональных компьютеров и недорогих устройств получения цифрового изображения

Одной из сравнительно новых областей применения алгоритмов распознавания является создание систем взаимодействия человека с компьютером. В общем случае технология взаимодействия основывается на получении видеоизображения с камеры и последующей обработке изображения в режиме реального времени.

Результатом обработки входных данных может являться дополнение изображения виртуальными элементами, то есть построение дополненной реальности [1]. Использование дополненной реальности в интернет-магазинах может увеличить их интерактивность и привлечь большее количество покупателей. Создание такого программного обеспечения позволит повысить интерактивность магазина и привлечь большее число покупателей.

Однако в настоящее время подобные приложения мало распространены. А существующие реализации имеют множество ограничений и недостатков. В данной статье рассматривается применение алгоритмов компьютерного зрения для создания виртуальной примерочной.

1 Анализ существующих реализаций

Существующие в настоящее время системы дополненной реальности, применяемые в качестве виртуальных примерочных, можно разделить на три категории:

- приложения, использующие изображения манекена;
- приложения, работающие с фотографией пользователя;
- приложения, работающие с видеоизображением, поступающим с камеры.

Примерами приложений, использующих первые два подхода, являются системы *odensya.com* и *VirtualDress*. В программном продукте *FashionIsta* используется изображение, полученное с USB-камеры, однако управление системой осуществляется при помощи жестов, положение корпуса человека при этом не анализируется.

На основе анализа существующих решений сформулированы следующие требования к разрабатываемой системе:

- приложение должно работать с видеоизображением, получаемым с USB-камеры;
- предмет одежды должен автоматически точно позиционироваться в зависимости от положения тела человека на изображении;
- изображение предмета одежды должно автоматически модифицироваться в зависимости от угла поворота тела и положения рук;
- алгоритм распознавания должен корректно работать при наличии постороннего шума на изображении, при различной освещенности и при неоднородном фоне.

2 Общий подход к решению задачи

На основе сформулированных требований можно предложить следующий алгоритм функционирования системы:

- 1) пользователю выдать предмет одежды с нанесенными метками (рисунок 1);
- 2) получить изображение с USB-камеры и произвести его предварительную обработку – уменьшить изображение, убрать шум, получить монохромное изображение;
- 3) распознать положение и ориентацию меток на изображении;
- 4) по найденным меткам определить ориентацию корпуса человека на изображении;
- 5) с учетом положения меток произвести пространственные преобразования модели предмета одежды (поворот, масштабирование) и наложить полученное изображение на исходное.



Рис. 1. Пример расположения меток

Для решения задачи необходимо использовать следующие алгоритмы обработки изображения: удаление шума, уменьшение изображения (дискретизация), пороговое отсечение и другие.

3 Обработка изображения

Для предварительной обработки изображения целесообразно использовать алгоритм свертки или линейных фильтров [2]. Простым примером свертки является вычисление среднего значения в фиксированной области. Учитываются все значения в пределах блока пикселей $(2k + 1) \times (2k + 1)$. Если на входе дано изображение F , то на выходе получится

$$R_{ij} = \frac{1}{(2k + 1)^2} \sum_{u=i-k}^{u=i+k} \sum_{v=j-k}^{v=j+k} F_{uv}.$$

Для удаления шума в данной работе использовалось симметричное гауссово ядро

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right),$$

где σ - среднее квадратичное отклонение гауссиана, измеряемое в расстояниях между пикселями.

Дискретное ядро сглаживания можно получить, построив массив $(2k + 1) \times (2k + 1)$, значения элемента (i, j) которого равно

$$H_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{((i - k - 1)^2 + (j - k - 1)^2)}{2\sigma^2}\right).$$

При малом σ ненулевым будет только один средний пиксель. При увеличении σ необходимо также увеличивать k , иначе не будет учтен вклад пикселей, которые должны входить значительными весовыми коэффициентами.

Для дальнейшей работы имеет смысл уменьшить размер изображения, то есть провести дискретизацию. При дискретизации часть информации теряется. Из теоремы Найквиста [3] следует, что сокращать изображение путем простой выборки каждого k -го пикселя может привести к сильным искажениям. Вместо этого можно фильтровать изображение таким образом, чтобы устранить пространственные частоты, превышающие новую частоту дискретизации.

Желаемого результата можно достичь, применив свертку с гауссовым ядром. Однако такой подход приводит к временным затратам на дополнительные вычисления.

Для определения целесообразности использования гауссовой свертки перед дискретизацией были подготовлены 50 различных изображений, полученных с USB-камеры с разрешением 640x480. При обработке изображений со сглаживанием и без него измерялась скорость работы алгоритмов и оценка выходного изображения.

Для вычисления оценки выходного изображения использовалось среднее отклонение значений яркости пикселей по шкале от 0 до 255 от сглаженного изображения.

В результате исследования алгоритмов на 50 примерах было получено абсолютное значение оценки выходного изображения 18 ± 2 единиц яркости, что соответствует относительной величине $0,045 \pm 0,001$. При этом дополнительное время на сглаживание составляло $0,0386 \pm 0,004$ с, что приблизительно равно 1/24 секунды.

Следовательно, использование сглаживания перед уменьшением изображения при решении поставленной задачи не приводит к большому искажению изображения, но требует дополнительных временных затрат.

Перед тем как приступить к распознаванию меток, необходимо провести отсечение по порогу яркости пикселей. Отсечение может быть проведено двумя способами: по глобальному значению порога для всего изображения или по локальному для каждой локальной области.

Для исследования первого способа в работе значение порога рассчитывалось как средняя величина яркости всех пикселей. Такой подход не требует больших вычислений, однако имеет существенный недостаток: при наличии «засвеченных» областей на изображении отсечение по порогу приводит к некорректному результату.

Второй подход заключается в разбиении исходного изображения на несколько областей, вычисление локального значения порога для каждой области и отсечения изображения в пределах этой области по рассчитанному значению локального порога.

Одновременно с вычисление локальных значений порога можно по гистограмме определить области изображения, заведомо не содержащие метки. Примеры гистограмм для разных областей показаны на рисунке 2. Видно, что областям, содержащим метки (рисунок 2, г), соответствуют гистограммы с двумя выраженными локальными максимумами. Значение порога отсечения следует выбрать равное середине интервала между двумя максимумами.

Полученные на выходе области изображения для определения положения и ориентации меток далее можно передать на вход обученной нейронной сети [4]. В данной работе для определения меток использовался поиск по шаблону с предварительной обработкой изображения градиентным фильтром.

Информация о найденных метках использовалась для модификации модели предмета одежды и получения итогового изображения.

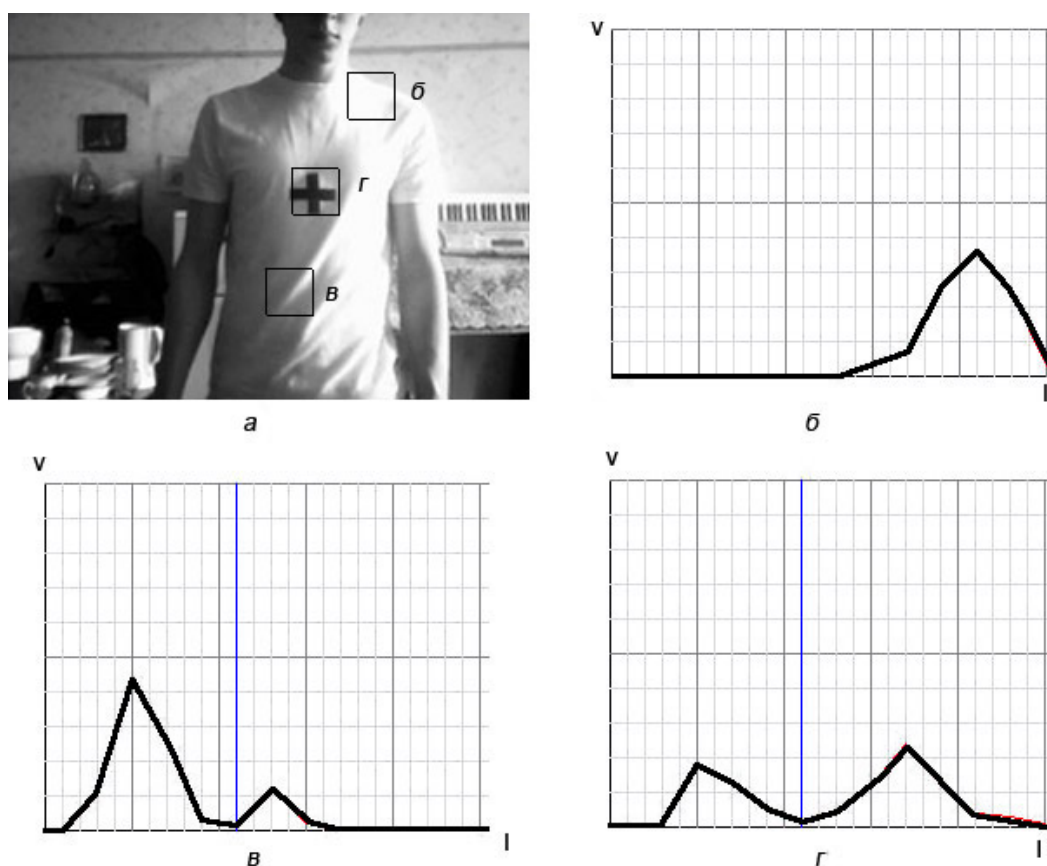


Рис. 2. Гистограммы, построенные для областей изображения

(I – яркость точки; v – частота): а – исходное изображение;

б, в – гистограммы, построенные по участкам, не содержащим метки; г – гистограмма, построенная для участка, содержащего фрагмент метки.

4 Заключение

В работе были проанализированы существующие системы, использующие технологии компьютерного зрения для построения дополненной реальности, на примере виртуальных примерочных. Выявлены их достоинства и недостатки. На основе проведенного анализа был сформулирован общий подход и предложен новый метод решения задачи.

Далее были сформулированы требования к разрабатываемому приложению и на их основе выбраны и исследованы алгоритмы компьютерного зрения, подходящие для решения задачи.

В ходе работы был разработан демонстрационный пример виртуальной примерочной. Также было проведено исследование предложенного метода, выявлена зависимость результата алгоритма от качества входного изображения, определены предельные параметры входных данных для функционирования приложения:

- контрастность изображения должна принадлежать отрезку $[0,4; 0,65]$;
- максимально допустимый угол поворота – 30 градусов;
- разрешение от 340x480.

Список литературы

- 1) Технология «Дополненная реальность» [Электронный ресурс]. – URL: <http://2nova.ru/lab/demoar.html>
- 2) Д. Форсайт, Ж. Понс Компьютерное зрение. Современный подход. – М.: Вильямс, 2004. – 928 с.
- 3) А.А. Харкевич Спектры и анализ. – М.: URSS, 2007. – 89 с.
- 4) Yin, Hongfeng (1996), Perceptron-Based Algorithms and Analysis, Spectrum Library, Concordia University, Canada