

Е.Н. Амиргалиев,
д.т.н. профессор,
Университет имени Сулеймана Демиреля
А.Н. Нусипбеков,
PhD –докторант,
КазНУ им. аль Фараби
Алматы/Казахстан

Распознавание жестов и технология реализации

Аннотация: В прикладных задачах распознавание жестов стало очень популярным в последние годы. Создано множество методов и технологий распознавания. Они очень часто применяются в различных областях науки и техники. В этой работе мы объясняем технологию реализации распознавания жестов на основе применения скрытых марковских моделей.

Ключевые слова: распознавание жестов, компьютерное зрение, марковская модель, жестовые языки, робототехника, устройство Kinect

Распознавание жестов это раздел в компьютерной науки, целью которого является интерпретация человеческих жестов с помощью различных методов. Жестами могут быть жесты рук, тела, лица. На данный момент популярны распознавание эмоций лица и движения рук. К настоящему времени на основе использования камеры и алгоритмов компьютерного зрения разработаны множество способов для интерпретирования жестовых языков.

1.Области применения

На сегодняшний день распознавание жестов применяется во многих областях, из которых можно выделить следующие:

- Жестовые языки – количество людей в мире с ограниченными возможностями общения велико, каждый восьмой человек в мире, к примеру, имеет проблемы со слухом. Распознавание жестов вносит свой огромный вклад для подобных людей, создавая различные технологии, дающие возможность общения.
- Робототехника – управление роботом с помощью жестов очень необходимо для роботов-помощников.
- Человеко-компьютерное взаимодействие – распознавание жестов, может быть рассмотрено как способ для компьютерной системы в понимании человеческого языка вместо обычных клавиатур и мышек (многомодальный ввод информации).
- Современные игровые системы – современные игровые приставки такие как Xbox.
- Медицина – реабилитационные машины.

2. Метод решения задачи

На сегодняшний день в основном используются различного рода акселерометры, датчики либо силуэты рук, полученные с обычных камер. Есть некоторые неудобства связанные с ними, такие как дороговизна некоторых технологий, зависимость от освещения и т.п.

Предлагаемый нами метод, это метод основанный на использовании устройства Kinect, который позволяет отслеживать руки при любом естественном освещении и является достаточно точным.

Kinect это устройство, выпущенное в конце 2011 года компанией Microsoft. Оно состоит из двух сенсоров глубины, цветной видеокамеры и микрофонной решетки. Датчик глубины состоит

из инфракрасного проектора, объединенного с монохромной матрицей, что позволяет датчику Kinect получать трёхмерное изображение при любом естественном освещении.

3. Реализация

Распознавание жестов имеет много общего со многими другими видами распознавания, в частности распознавание речи, биометрики: т.е. имея сигнал меняющийся во времени, как мы можем узнать насколько он соответствует имеющейся модели. Обычно для подобных видов задач используются скрытые модели маркова. Жест руки – это последовательность множества точек в 3D пространстве (Рис. 1).

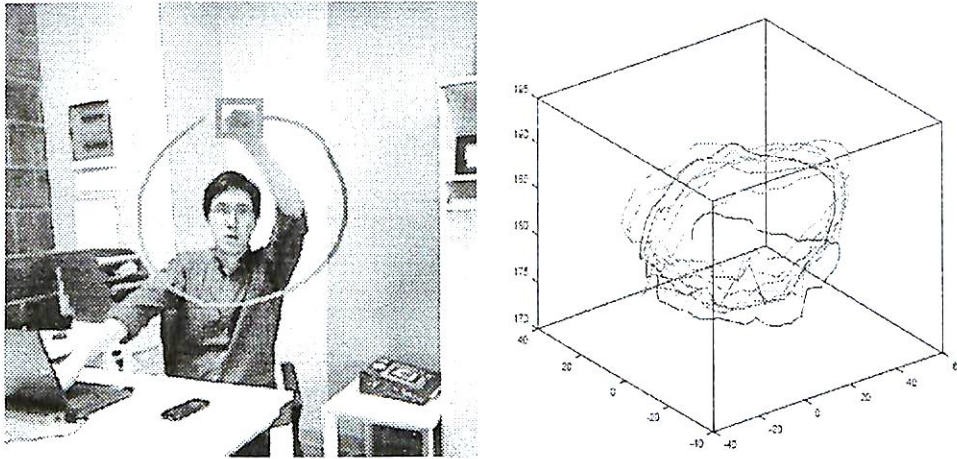


Рисунок 1- Круговой жест руки

Жест может быть представлен как Марковская цепь, где истинные состояния $S = \{s_1, s_2, s_3, \dots, s_N\}$ скрыты. А наблюдаться может $O = \{o_1, o_2, \dots, o_M\}$ (позиция руки), которая распределяется с определенной вероятностью от каждого состояния.

Мы вычисляем значения вероятностей переходов между состояниями A и вероятности результатов B во время «обучения» с помощью алгоритма Баума Велша, который итеративно вычисляет параметр $\lambda = (A, B)$. При заданном векторе наблюдений O алгоритм находит

$$\lambda^* = \max_{\lambda} P(O | \lambda)$$

Элементы матрицы переходов вычисляются по следующей формуле:

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\hat{b}_j(o_k) = \frac{\sum_{t: o_t = o_k} \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}$$

Также нам нужно будет воспользоваться алгоритмом «вперед-назад».

Сначала вычисляется

$$\hat{\alpha}_i(t) = p(O_1 = o_1, \dots, O_t = o_t, Q_t = i | \lambda),$$

что является вероятностью получения заданной последовательности O_1, \dots, O_t для состояния i в момент времени t . Она вычисляется рекурсивно:

$$\alpha_j(t+1) = b_j(O_{t+1}) \sum_{i=1}^N \alpha_i(t) \cdot a_{ij}.$$

Обратная процедура позволяет вычислить вероятность конечной заданной последовательности при условии, что мы начали из исходного состояния, в момент времени t :

$$\beta_i(t) = \sum_{j=1}^N \beta_j(t+1) a_{ij} b_j(O_{t+1}).$$

Используя α и β можно вычислить величину вероятности скрытого состояния i в момент t и вероятность перехода из i в j :

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{Pr(O|\lambda)}$$

$$\gamma_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{Pr(O|\lambda)}$$

Заключение

Для реализации были написаны приложения на C#(Visual Studio) для отслеживания руки с использованием Kinect SDK. А также использованы вышеупомянутые алгоритмы реализованные на Octave. Были использованы жесты цифр в воздухе. При тестировании данный способ, показал хорошие результаты, например, в среднем для несложных жестов, такие как круг, результаты были выше 90%. В целом можно заключить, что данный метод дает возможность определять жесты руки с очень хорошей точностью.

Литература

- 1 Рабинер Л.Р, «Скрытые марковские модели и их применение в избранных приложениях при распознавании речи», IEEE, 77, No. 2, 1989.
- 2 Николенко С. Лекции № 6 и № 7 (слайды), посвящённые скрытым марковским моделям, из курса «Вероятностное обучение»
- 3 Джосан О.В. Использование Скрытых Марковских Моделей для детектирования радужки на изображении лица, Лекции Московского Государственного Университета им. М.В. Ломоносова, Москва, Россия
- 4 Николенко С. Лекции № 6 и № 7 (слайды), посвящённые скрытым марковским моделям, из курса «Вероятностное обучение». МГУ имени М.В. Ломоносова
- 5 Нандо де Фрейтас, «Скрытые марковские модели». Лекции CPSC340, университет Британской Колумбии, сентябрь 2012
- 5 Ямато, Юнжи, «Распознавание человеческих жестов используя скрытые модели маркова». Конференция IEEE. Япония 15-18 июня 1992.

Түйін

Макалада сахына танудың кажетті косымша есептері ретінде карастырылатын кол кимылын тану есебі карастырылған. Осы есептің шешімі жасырын марков үдерістеріне негізделген.

Özet

Makalede genellikle bir bileşen sahne tanıma olarak kullanılan el hareketleri, tanıma sorunu incelenmektedir. Ve bu araştırmadaki esas probleminin çözümü olarak Gizli Markov süreçleri kullanılmıştır.

Resume

The paper considers the problem of recognizing hand gestures, which in turn is often used as a component scene recognition. Solution of the problem based on hidden Markov processes.