

**ОТСЛЕЖИВАНИЕ ДВИЖЕНИЙ ПАЛЬЦЕВ
В ИЗОБРАЖЕНИЯХ КАМЕРЫ КРУГОВОГО ОБЗОРА
ПОСРЕДСТВОМ АЛГОРИТМА CONDENSATION ***

Н. Икома¹, М. Саката², М. Дои³

¹*Kyushu Institute of Technology, Fukuoka, Japan
E-mail: ikoma@comp.kyutech.ac.jp*

²*Hitachi Kokusai Electric Inc., Tokyo, Japan
E-mail: sakata@sys2.comp.kyutech.ac.jp*

³*Osaka Electro-Communication University, Osaka, Japan
E-mail: doi@isc.osakac.ac.jp*

Исследуется возможность слежения за изображениями пальцев в камере кругового обзора с целью создания дружелюбного человеку интерфейса. Для его разработки необходимо знать положение пальцев на изображении. Предложен новый метод слежения за пальцами, использующий алгоритм CONDENSATION (Condition Density Propagation), который является специальным случаем фильтра частиц (particle filters). Разработанный метод основан на моделировании движений пальцев в пространстве состояний, где вектор состояния модели формируется из состояний каждого из всех отслеживаемых пальцев. Экспериментальные результаты демонстрируют работоспособность предложенного метода.

Введение. Создание дружелюбного человеку интерфейса является важной задачей в разработке компьютерных систем. Его назначение – обеспечить пользователям любого уровня легкий способ взаимодействия с системой. Жестикуляция является ключом к достижению такого взаимодействия при помощи систем технического зрения, с недавних пор сочетающих низкую цену с высокой эффективностью.

Среди различных подходов, использующих для такого взаимодействия движение пальцев, представляет интерес применение камеры кругового обзора [1]. Камера используется необычным способом, при котором пользователь захватывает ее рукой, как показано на рис. 1. Здесь присутствует несколько типичных движений руки и пальцев, таких как вращение запястья, продольное скольжение руки, движение каждого пальца и перехватывание пальцами. Типичные изображения, полученные при этом, показаны на рис. 2. Каждое движение пальцев на изображении обусловлено движением

* Работа выполнена при поддержке ISM Cooperative Research Program (2004-ISM-CRP-2020) и частично при поддержке Grant-in-Aid for Science Research (A) (14208025).

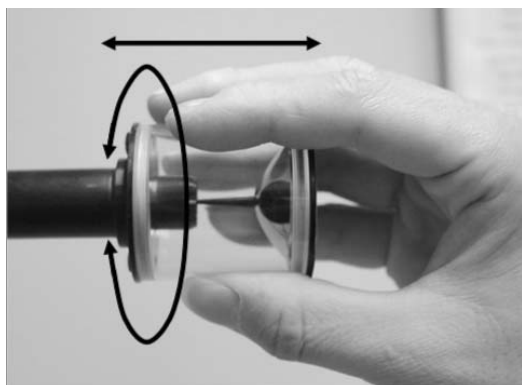


Рис. 1. Изображения пальцев, охватывающих камеру кругового обзора. Типичные движения: вращение и скольжение пальцев (показаны стрелками)

руки, например вращение всех пальцев вызывается вращением запястья, изменение длины пальцев – продольным движением руки, небольшие перемещения каждого пальца – движением пальцев в отдельности, а появление и исчезновение пальцев – перехватыванием.

Чтобы создать дружелюбный интерфейс, который использует изображения пальцев, сформированные камерой кругового обзора, необходимо знать состояние пальцев: положение, угол наклона, длину и ширину пальца, а также скорость их движения. Предлагается метод оценивания состояния пальцев посредством моделирования пространства состояний. Состояние модели – это множество состояний всех пальцев, подлежащих отслеживанию. Модель пространства состояний включает в себя модель системы, которая представляет собой движение пальцев, и модель наблюдений, которая применяется для вычисления правдоподобия. В качестве модели системы используется случайное блуждание состояния в предположении гладкости движения и малой неопределенности. Из модели наблюдений посредством

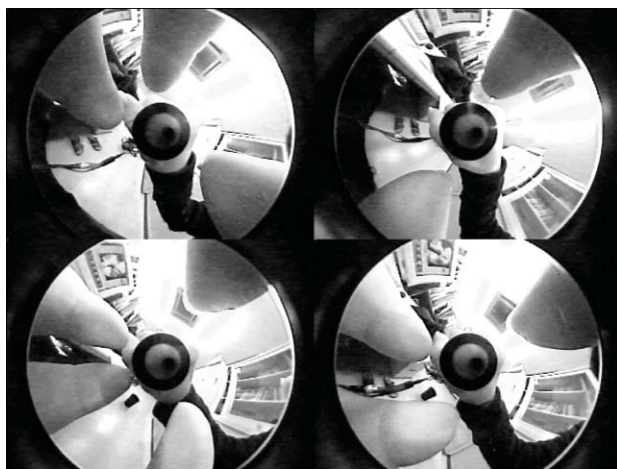


Рис. 2. Типичные положения пальцев

сравнения состояния с реальным изображением контуров пальцев находится правдоподобие.

Идея использования информации о контуре известного объекта, подлежащего отслеживанию, восходит к алгоритму CONDENSATION (Condition Density Propagation) [2]. Это специальный случай фильтра частиц [3], являющегося методом оценивания состояния посредством аппроксимации состояния множеством частиц в пространстве состояний. Алгоритм CONDENSATION предполагает, что типичная форма отслеживаемого объекта известна и хранится в виде шаблона. Для отслеживания в изображении используется аффинное преобразование шаблона. В рамках фильтра частиц каждая частица представляет собой гипотетическое аффинное преобразование. На протяжении трех этапов, состоящих из предсказания, наблюдения и селекции, множество частиц аппроксимирует апостериорное распределение состояния, заданного кадрами изображения до текущего момента, т. е. оценки состояния. На этапе предсказания каждая частица смещается в соответствии с моделью системы. Функция правдоподобия каждой частицы вычисляется на этапе наблюдения. А на этапе селекции с вероятностью, пропорциональной правдоподобию, выполняется выбор с возвращением. Затем по множеству частиц мы получаем оценку состояния пальцев.

Данная работа построена следующим образом. Разд. 1 содержит описание проблемы отслеживания пальцев в изображениях камеры кругового обзора. В математической модели (разд. 2) проблема сформулирована в предположении, что количество пальцев известно и фиксировано. Метод оценивания состояния, использующий алгоритм CONDENSATION, рассматривается в разд. 3. Экспериментальные результаты по отслеживанию движений одного, двух, трех и четырех пальцев приводятся в разд. 4.

1. Постановка задачи. Задача состоит в создании дружественного человеку интерфейса, использующего изображения пальцев, полученные камерой кругового обзора в режиме, показанном на рис. 1. Необходимо выделить информацию о движении пальцев из типичной последовательности изображений (рис. 2). Первым шагом в достижении этой цели является разработка метода отслеживания для известного и фиксированного количества пальцев. Допускаются любые движения пальцев, охватывающих цилиндр камеры, пока все отслеживаемые пальцы наблюдаются на изображении. При этом порядок пальцев не изменяется и перехватывание не допускается.

Дружественный интерфейс предполагает работу в реальном времени. В связи с этим при отслеживании используется только текущий кадр и никогда не запоминаются предыдущие, поэтому обновление отслеживаемой информации основано только на текущем кадре изображения. Движение пальцев является плавным в последовательности кадров, поскольку частота кадров достаточно высокая. Камера работает в обычных условиях, например в комнате без специальных средств подавления фоновых объектов на изображении. Следовательно, на изображениях кроме пальцев присутствуют другие подвижные и неподвижные объекты.

2. Математическая модель. 2.1. *Состояние пальца.* Представление пальца на изображении в камере кругового обзора может быть упрощено, что иллюстрирует рис. 3. Палец можно представить центральным углом θ , шириной w , высотой h и наклоном ϕ .

Это упрощение используется следующим образом. Пальцы имеют типичную форму, и каждое изображение пальца идентифицируется положени-

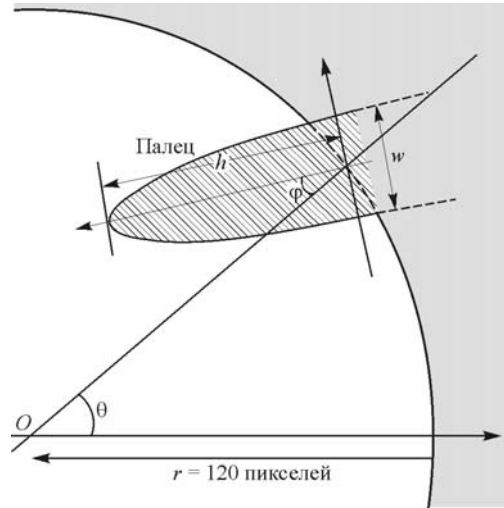


Рис. 3. Моделирование пальца в изображении, полученном камерой кругового обзора (O – центр изображения)

ем θ , шириной w , высотой h и наклоном φ . Соответственно набор этих параметров формирует состояние пальца

$$\mathbf{x} = [\theta, w, h, \varphi]. \quad (1)$$

Назовем \mathbf{x} вектором состояния пальца.

Введем для вектора состояния индекс пальца $f = 1, 2, \dots, F$ (F – известное и фиксированное количество пальцев) и индекс дискретного времени $k = 1, 2, \dots$, обозначив соответствующий вектор как $\mathbf{x}_{f, k}$. Эти индексы применимы и к элементам вектора состояния.

2.2. *Модель системы.* Фундаментальная идея модели системы в рамках моделирования пространства состояний заключается в использовании модели случайного блуждания для временной эволюции вектора состояния каждого пальца. Модель случайного блуждания записывается в виде стохастического дифференциального уравнения

$$\mathbf{x}_{f, k} = \mathbf{x}_{f, k-1} + \mathbf{v}_{f, k}, \quad (2)$$

где $\mathbf{v}_{f, k}$ – случайный вектор, называемый «шумом системы», который играет роль неопределенности в движении пальца. Обычно шум системы предполагают подчиняющимся распределению Гаусса с нулевым средним и диагональной ковариационной матрицей. Альтернативное представление модели задается условной плотностью вероятности

$$\mathbf{x}_{f, k} \sim f(\mathbf{x} | \mathbf{x}_{f, k-1}). \quad (3)$$

Реализация модели системы включает детальные условия, такие как допустимый диапазон значений каждой переменной в векторе состояния. Чтобы не нарушить диапазон каждой переменной, выполняется соответствующее

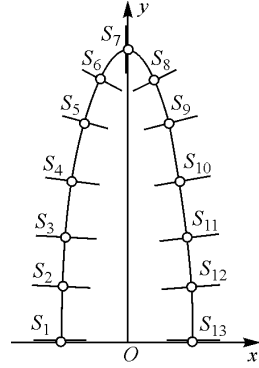


Рис. 4. Шаблон изображения пальца. Кривой обозначен типичный контур пальца, кружками – контрольные точки, а каждая линия, пересекающая кружок, обозначает отрезок прямой

ющее преобразование, например логарифм. Он используется для положительных параметров, таких как ширина w и высота h . Необходимо также сохранять порядок пальцев при преобразовании параметров центрального угла $\theta_f, f = 1, 2, \dots, F$, к $\varphi_1 = \log \theta_1, \varphi_2 = \log(\theta_2 - \theta_1), \dots, \varphi_F = \log(\theta_F - \theta_{F-1})$ по модулю 2π .

Состояние всех пальцев представлено совокупностью индивидуальных векторов состояния:

$$\mathbf{X}_k = [\mathbf{x}_{1,k}, \mathbf{x}_{2,k}, \dots, \mathbf{x}_{F,k}]. \quad (4)$$

Эволюция всех пальцев во времени дана в форме стохастического дифференциального уравнения модели случайного блуждания

$$\mathbf{X}_k = \mathbf{X}_{k-1} + \mathbf{V}_k \quad (5)$$

с $\mathbf{V}_k = [\mathbf{v}_{1,k}, \mathbf{v}_{2,k}, \dots, \mathbf{v}_{F,k}]$ или в альтернативной форме условной плотности вероятности

$$\mathbf{X}_k \sim f(\mathbf{X} | \mathbf{X}_{k-1}). \quad (6)$$

2.3. Функция правдоподобия. В соответствии с идеей алгоритма CONDENSATION [2] типичная форма отслеживаемого объекта запоминается в качестве шаблона. В нашем методе шаблоном является типичная форма пальца, показанная на рис. 4. В определенных местах контура пальца на шаблоне размещены $P = 13$ контрольных точек, и через каждую контрольную точку проходит отрезок прямой длиной l для измерений на изображении.

Шаблон преобразуется в соответствии с вектором состояний $\mathbf{x}_{f,k}$, который запоминает параметры аффинного преобразования f -го пальца в момент k . Пусть \mathbf{I}_k – кадр изображения в момент k . Аффинно-преобразованный шаблон накладывается на изображение. Затем получают функции интенсивности изображения вдоль каждого отрезка прямой преобразованного шаблона в виде

$$I_{j,k}(z; \mathbf{x}_{f,k}) \quad (7)$$

для j -й контрольной точки f -го пальца, где z – его положение на отрезке прямой с началом в контрольной точке и осью, направленной наружу из области пальца.

Вычисление первой разности от функции интенсивности (7) дает $|I'_{j,k}(z; \mathbf{x}_{f,k})|$. Наблюдением на каждом отрезке прямой является положение ближайшего к началу значения $|I'_{j,k}(z; \mathbf{x}_{f,k})|$, превышающего порог. Набор наблюдений для f -го пальца в момент k обозначим как

$$\mathbf{z}_{f,k} = [z_{1,f,k}, z_{2,f,k}, \dots, z_{P,f,k}]. \quad (8)$$

Модель наблюдений представляется плотностью вероятности наблюдений $p(z)$ вдоль каждого линейного сегмента. Предположим, что плотности $p(z)$ взаимно независимы, тогда функция правдоподобия f -го пальца принимает вид

$$h(\mathbf{z}_{f,k} | \mathbf{x}_{f,k}) \equiv \prod_{j=1}^P p(z_{j,f,k}). \quad (9)$$

Типичной формой плотности наблюдений является усеченное распределение Гаусса

$$p(z; \sigma^2, l) \propto \sigma^{-1} \exp \{-\min(z^2, (l/2)^2)/2\sigma^2\} \quad (10)$$

с дисперсией шума наблюдений σ^2 и длиной сегмента l .

Функция правдоподобия для всех пальцев в предположении взаимной независимости пальцев есть произведение функций правдоподобия каждого пальца:

$$h(\mathbf{I}_k | \mathbf{X}_k) \equiv h(\mathbf{Z}_k | \mathbf{X}_k) = \prod_{f=1}^F h(\mathbf{z}_{f,k} | \mathbf{x}_{f,k}), \quad (11)$$

где $\mathbf{Z}_k = [\mathbf{z}_{1,k}, \mathbf{z}_{2,k}, \dots, \mathbf{z}_{F,k}]$ – все наблюдения в момент k .

3. Оценивание положения пальцев заключается в получении апостериорного распределения состояния, заданного всеми кадрами изображения до текущего момента времени. Поскольку согласно требованиям задачи доступны только текущий кадр изображения, а все предыдущие кадры должны отбрасываться, используется процедура последовательного оценивания, которое должно обновлять апостериорное распределение следующим образом:

$$p(\mathbf{X}_{1:k} | \mathbf{I}_{1:k}) \propto p(\mathbf{X}_{1:k-1} | \mathbf{I}_{1:k-1}) h(\mathbf{I}_k | \mathbf{X}_k) f(\mathbf{X}_k | \mathbf{X}_{k-1}). \quad (12)$$

Здесь $\mathbf{I}_{1:k} \equiv (\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_k)$. Заметим, что апостериорное распределение в терминах $\mathbf{X}_{1:k}$ использовано в (12) потому, что уравнение обновления записывается таким образом проще, чем с использованием \mathbf{X}_k . Это только вывод математической формулировки. В действительности вычисление выполняется для самого последнего индекса времени при \mathbf{X} , т. е. $p(\mathbf{X}_k | \mathbf{I}_{1:k})$ находится из $p(\mathbf{X}_{k-1} | \mathbf{X}_{1:k-1})$.

В общем случае уравнение (12) не может быть решено в замкнутом виде. Такое решение возможно только для небольшого числа моделей, например для линейной гауссовой модели пространства состояний. В этом случае апостериорное распределение становится гауссовым, и решение получается с помощью фильтра Калмана посредством вычисления вектора средних значений и ковариационной матрицы распределения. Вышепредставленная модель к таковым не относится, поэтому решать уравнение следует альтернативным методом.

Идея фильтра частиц [3] заключается в получении приближенного решения уравнения (12) посредством использования множества частиц в пространстве состояний. Алгоритм CONDENSATION [2] является одной из типичных реализаций фильтра частиц, которая использует шаблон и аффинное преобразование, чтобы вычислить правдоподобие для визуального отслежи-

вания. Бутстрэп-фильтр [4] и фильтр Монте-Карло [5] также являются специальными реализациями фильтра частиц. Более того, само понятие фильтра частиц является специальным случаем последовательного метода Монте-Карло [3] в контексте оптимальной фильтрации. Распределение, оцениваемое посредством последовательного метода Монте-Карло, не обязательно должно быть апостериорным распределением (12) [3].

Начиная с исходного множества частиц, соответствующего исходному распределению, алгоритм CONDENSATION обновляет его согласно уравнению (12). Далее алгоритм представим детально.

3.1. *Инициализация.* Пусть $\{\mathbf{X}_0^{(i)}\}_{i=1}^M$ – исходное множество, состоящее из M частиц. Исходные частицы генерируются согласно данному исходному распределению $p(\mathbf{X}_0 | \mathbf{I}_{0:0})$.

3.2. *Итерации.* Начиная с исходного множества частиц и полагая $k := 1$, обновление выполняем в соответствии с (12) в три этапа (предсказание, наблюдение и селекция) с последующим наращиванием индекса времени $k := k + 1$.

3.2.1. *Этап предсказания.* Переместим частицы для текущего момента k с помощью

$$\tilde{\mathbf{X}}_k^{(i)} \sim f(\mathbf{X} | \mathbf{X}_{k-1}^{(i)}), \quad i=1,2,\dots,M. \quad (13)$$

Полученное множество частиц $\{\mathbf{X}_k^{(i)}\}_{i=1}^M$ аппроксимирует предсказанное на один шаг распределение $p(\mathbf{X}_k | \mathbf{X}_{1:k-1})$.

3.2.2. *Этап наблюдения.* Вычисляется правдоподобие для каждой частицы и вес

$$\alpha_k^{(i)} \propto h(\mathbf{I}_k | \tilde{\mathbf{X}}_k^{(i)}), \quad i=1,2,\dots,M, \quad (14)$$

где веса нормированы так, чтобы $\sum_{i=1}^M \alpha_k^{(i)} = 1$. Полученное таким образом множество частиц с весами $\{(\tilde{\mathbf{X}}_k^{(i)}, \alpha_k^{(i)})\}_{i=1}^M$ аппроксимирует апостериорное распределение $p(\mathbf{X}_k | \mathbf{I}_{1:k})$.

3.2.3. *Этап селекции.* На этом этапе проводится выборка с возвращением для взвешенных частиц $\{(\tilde{\mathbf{X}}_k^{(i)}, \alpha_k^{(i)})\}_{i=1}^M$. Выбирается M частиц с вероятностью, пропорциональной весу, т. е. выполняется M повторений выборок с возвращением с вероятностью $\alpha_k^{(i)}$ частицы $\tilde{\mathbf{X}}_k^{(i)}$ для $i=1,2,\dots,M$. Запишем формально эту процедуру в виде

$$\mathbf{X}_k^{(i)} = \begin{cases} \tilde{\mathbf{X}}_k^{(1)} & \text{с вероятностью } \alpha_k^{(1)}; \\ \tilde{\mathbf{X}}_k^{(2)} & \text{с вероятностью } \alpha_k^{(2)}; \\ \vdots & \\ \tilde{\mathbf{X}}_k^{(M)} & \text{с вероятностью } \alpha_k^{(M)} \end{cases} \quad (15)$$

для $i=1,2,\dots,M$. Полученное множество из M частиц $\{\mathbf{X}_k^{(i)}\}_{i=1}^M$ тоже аппроксимирует апостериорное распределение $p(\mathbf{X}_k | \mathbf{I}_{1:k})$.

3.3. *Вычисление оценки.* Чтобы получить результат отслеживания, необходимо вычислить значение оценки по множеству частиц, которые аппрок-

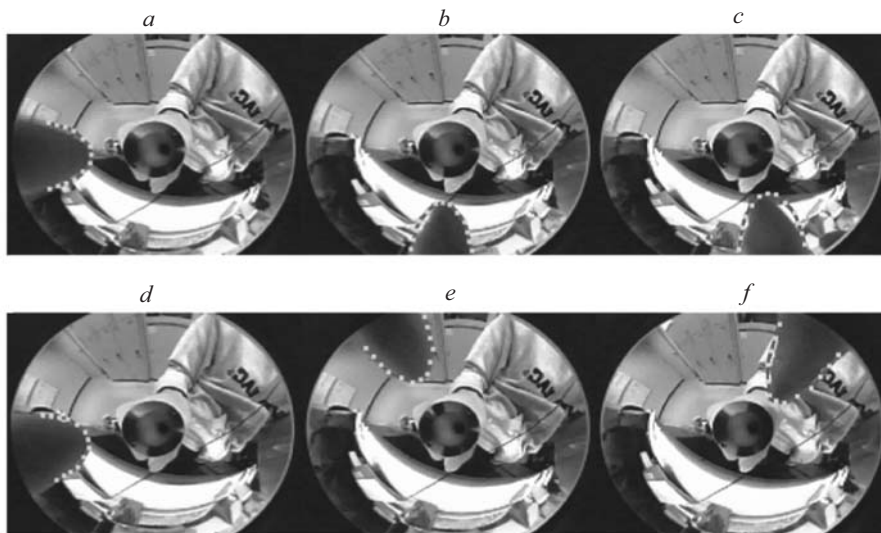


Рис. 5. Результат слежения за одним пальцем в последовательности, состоящей из 300 кадров: 50-й кадр (a), 100-й кадр (b), 150-й кадр (c), 200-й кадр (d), 250-й кадр (e), 300-й кадр (f)

симируют апостериорное распределение. Для этого используется среднее значение распределения

$$\bar{\mathbf{X}}_k = \frac{1}{M} \sum_{i=1}^M \mathbf{X}_k^{(i)}. \quad (16)$$

4. Эксперименты. В данной работе представлено четыре эксперимента с числом пальцев $F = 1, 2, 3, 4$, чтобы показать действие предложенного метода.

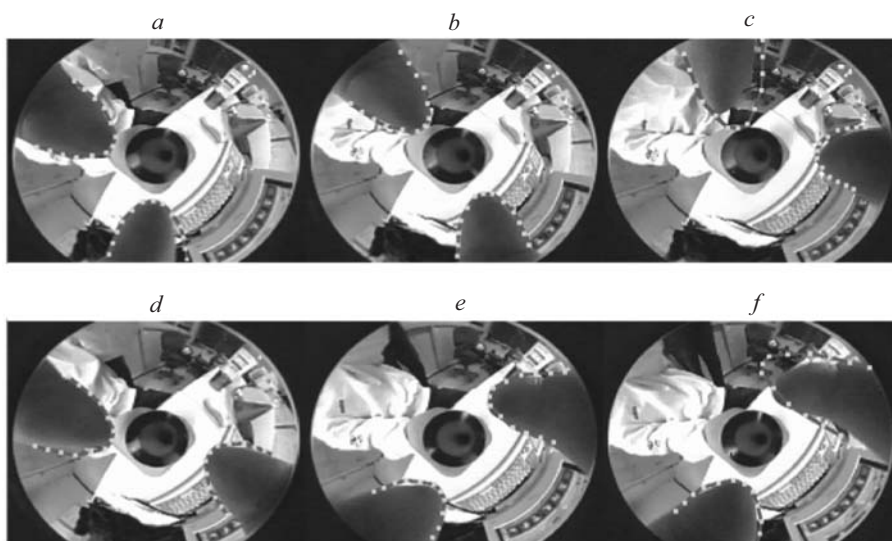


Рис. 6. Результат слежения за двумя пальцами в последовательности, состоящей из 300 кадров (последовательность кадров такая же, как на рис. 5)

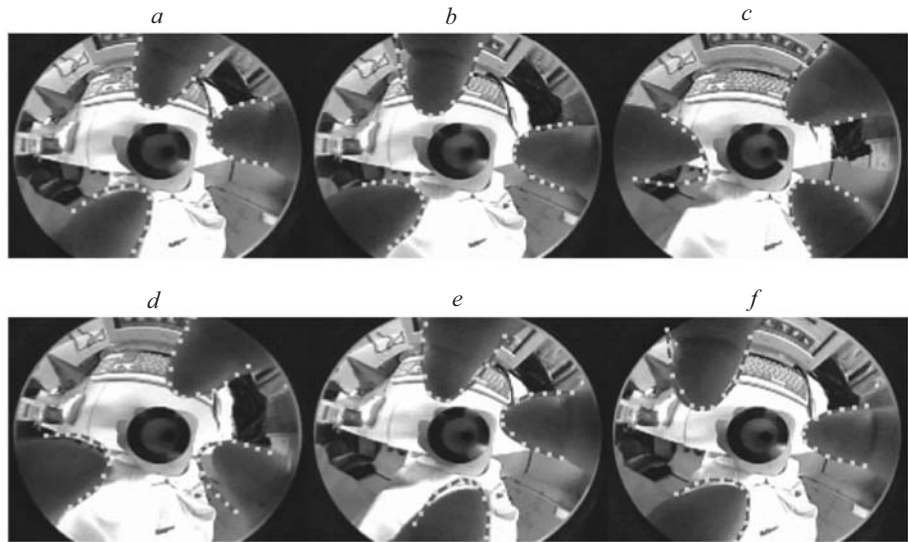


Рис. 7. Результат слежения за тремя пальцами в последовательности, состоящей из 300 кадров (последовательность кадров такая же, как на рис. 5)

Камера кругового обзора удерживалась левой рукой, а правая рука использовалась при работе с интерфейсом. Было получено четыре последовательности кадров для $F = 1, 2, 3, 4$ пальцев.

Условия оценивания общие для всех экспериментов и состоят в следующем. Количество частиц $M = 10000$. Дисперсия шума наблюдения $\sigma^2 = 16$ при длине линейного сегмента в шаблоне $l = 3\sigma/2$. Системный шум предполагается подчиняющимся взаимно независимым нормальным распределениям с дисперсиями $\tau_\theta^2 = 0,07 \text{ рад}^2$, $\tau_w^2 = 0,03$, $\tau_h^2 = 3,00$ и $\tau_\varphi^2 = 0,03 \text{ рад}^2$. Началь-

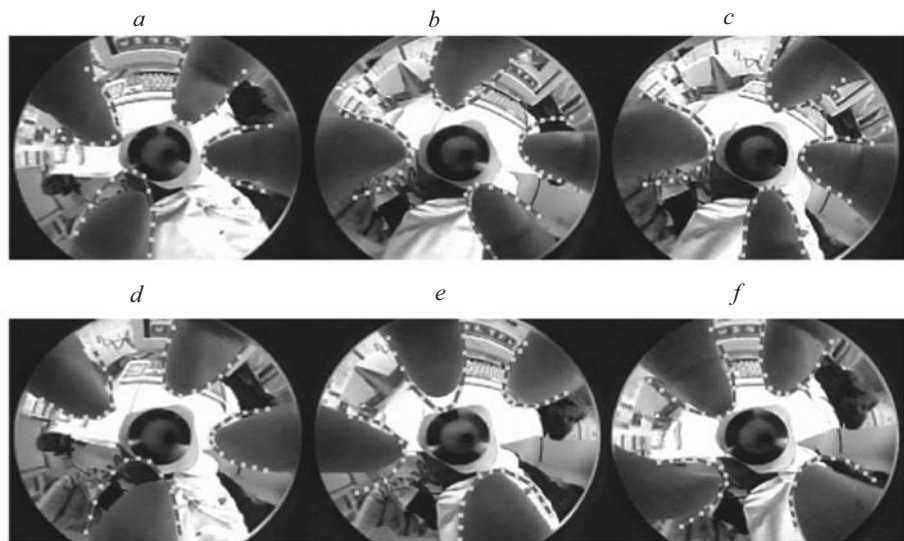


Рис. 8. Результат слежения за четырьмя пальцами в последовательности, состоящей из 300 кадров (последовательность кадров такая же, как на рис. 5)

ными распределениями служат следующие. Для θ используется $N(\bar{\theta}, 0,07)$, причем $\bar{\theta}$ получается вручную из начального изображения; для w и h – равномерные распределения $U[w_{\min}, w_{\max}]$ и $U[h_{\min}, h_{\max}]$, где $w_{\max} = 20$, $w_{\min} = 40$, $h_{\min} = 10$, $h_{\max} = 70$; для $\varphi - N(0, 0,03)$.

Результаты отслеживания движений пальцев показаны на рис. 5–8 для $F = 1, 2, 3, 4$ соответственно. На этих рисунках приведены некоторые характерные изображения. Поверх каждого исходного изображения нанесены контрольные точки и шаблоны. Они вычислены по оцененным в соответствии с (16) величинам. Исходя из рисунков можно сделать вывод, что полученные приемлемые результаты отслеживания.

Заключение. В данной работе предложен метод визуального отслеживания пальцев в изображениях камеры кругового обзора, который предназначен для нового дружественного человеку интерфейса, использующего движения пальцев, а также модель пространства состояний с плавным движением пальцев. В этой модели для процесса наблюдения используется шаблон пальца, а для представления каждого пальца в изображении применяется аффинное преобразование шаблона. Это соответствует идее алгоритма CONDENSATION [2].

Состояние пальцев оценивается с применением фильтра частиц, и получается множество частиц, которое аппроксимирует апостериорное распределение вектора состояния в текущий момент времени, определяемого всеми кадрами вплоть до текущего момента. Последовательно применяя эту процедуру оценивания, мы получили набор результатов оценивания, которые формируют траекторию последовательных движений пальцев. Экспериментальные результаты показали работоспособность предложенного метода.

Данная модель разработана для известного и фиксированного количества пальцев. В перспективе мы модифицируем метод для работы с переменным количеством пальцев. Для разработки интерфейса, использующего движение пальцев, необходимо обеспечить работу алгоритма в реальном времени.

СПИСОК ЛИТЕРАТУРЫ

1. **Doi M., Ueda S., Akiyama K.** Human interface based on finger gesture recognition using omnidirectional image sensor // Proc. of the 2003 IEEE Intern. Symp. on Virtual Environments, Human-Computer Interfaces, and Measurement Systems (VECIMS 2003). Lugano, Switzerland, 2003. P. 68.
2. **Isard M., Blake A.** CONDENSATION – Conditional density propagation for visual tracking // Intern. Journ. Computer Vision. 1998. **29**, N 1. P. 5.
3. **Monte Carlo** Methods in Practice /Eds. A. Doucet, J. F. G. de Freitas, N. J. Gordon. N. Y.: Springer, 2001.
4. **Gordon N. J., Salmond D. J., Smith A. F. M.** Novel approach to nonlinear / non-Gaussian Bayesian state estimation // IEE Proc. F. 1993. **140**, N 2. P. 107.
5. **Kitagawa G.** Monte Carlo filter and smoother for non-Gaussian nonlinear state space models // Journ. Computat. and Graphical Statist. 1996. **5**, N 1. P. 1.

Поступила в редакцию 30 сентября 2005 г.