

УДК 621.317.08 (075.8) © В.А. КНЯЗЬ, И.В. КРУПЕННИКОВ, А.В. ШАРОНОВ, 2008

ВИДЕОИЗМЕРЕНИЯ И ИХ КОРРЕЛЯЦИОННОЕ СТЕРЕООТОЖДЕСТВЛЕНИЕ, ИСПОЛЬЗУЮЩЕЕ ГРАДИЕНТНЫЙ ПОДХОД И ПИРАМИДЫ СТЕРЕОПАР

Владимир Александрович КНЯЗЬ родился в 1957 г. в городе Калининграде Московской области. Начальник лаборатории ФГУП «ГосНИИАС». Основные научные интересы — в области создания алгоритмического и программного обеспечения для компьютерных систем анализа и обработки изображений, а также ближней фотограмметрии. Автор более 85 научных работ.

Vladimir A. KNYAZ, was born in 1957, in Moscow Region. He is the Head of a Laboratory at the State Research Institute of Aviation Systems, Russian State Scientific Center (FGUP GosNIIAS). His research interests are in the algorithms and software for computer-aided systems aimed to analysis and processing of images as well as in the short-range photogrammetry. He has published over 85 technical papers.

Илья Владимирович КРУПЕННИКОВ родился в 1983 г. в городе Москве. Аспирант МАИ. Основные научные интересы — в области трехмерной визуализации и технического зрения. Автор восьми научных работ.

Ilya V. KRUPENNIKOV, was born in 1983, in Moscow. He is Post-Graduate Student at the MAI. His research interests are in the three-dimensional visualization and computer vision. He has published 8 technical papers.

Анатолий Васильевич ШАРОНОВ родился в 1937 г. в городе Москве. Заведующий кафедрой МАИ. Доктор технических наук, профессор. Основные научные интересы — в области математической обработки результатов измерений. Автор более 150 научных работ.

Anatoly V. SHARONOV, D.Sci., was born in 1937, in Moscow. He is the Head of a Department at the MAI. His research interests are in the mathematical treatment of measurement data. He has published over 150 technical papers.

В задачах видеоизмерений для целей управления движением динамического объекта, использующих видеоинформацию со стереопар, критическим параметром является время стереотождествления. Поэтому для быстрого стереотождествления точек снимков, полученных с двух камер с известными параметрами взаимного ориентирования, предлагается алгоритм поиска соответствующих точек, использующий градиентный подход и пирамиду изображений. Алгоритм реализован на экспериментальной установке с цифровыми ПЗС-камерами с разрешением 640x480 пикселей и частотой съемки 30 кадров в секунду.

Известны работы [1, 2] в которых решаются задачи использования видеоизмерений, получаемых с двух ПЗС-камер, для разработки корреляционно-го алгоритма стереотождествления точек левого и правого снимков. В статье предлагается метод решения задачи стереотождествления, основанный на анализе реальных изображений подстилающей

поверхности с подвижного объекта. При этом учитываются основные факторы, затрудняющие стереотождествление:

- существенные яркостные различия изображений стереопары, возникающие при съемке поверхностей объектов под разными углами;

- существенные геометрические искажения вследствие потенциально сложного рельефа подстилающей поверхности;
- возможность наличия на изображениях областей с малой вариацией яркости.

Точность привязки к эталонному изображению тем выше, чем более неоднородным являются соответствующие фрагменты изображения. С вычислительной точки зрения самым простым и эффективным показателем неоднородности фрагмента является дисперсия яркости внутри фрагмента

$$\sigma^2(x_0, y_0, N) = \frac{1}{(2N+1)^2} \sum_{x=-N}^N \sum_{y=-N}^N (f(x+x_0, y+y_0))^2 - \left(\frac{1}{(2N+1)^2} \sum_{x=-N}^N \sum_{y=-N}^N (f(x+x_0, y+y_0)) \right)^2,$$

где $f(x+x_0, y+y_0)$ — функция яркости изображения относительно эталона; N — количество пикселей в изображении; x_0, y_0 — координаты эталонного изображения.

Чем выше дисперсия, тем более неоднородным является фрагмент.

Пусть изображения имеют размеры $M \times N$ пикселей и область перекрытия составляет процентов. Тогда для эталона функции яркости f левого изображения в области перекрытия правого изображения должен быть произведен поиск образа f . Это требует порядка $MN_n \alpha$ операций, где N_n — количество пикселей в эталоне. Если необходимо найти соответствие всем пикселям левого изображения, требуется $k = (MN_n \alpha)^2$ операций. Для реальных стереопар $\alpha \approx 0,6$, $M > 1000$, поэтому полное время стереотождествления становится недопустимо большим.

В процессе изучения снимков подстилающей поверхности, полученных ПЗС-камерами, установленными на движущемся объекте, выяснилось, что использование яркостных признаков для отождествления может привести к ненадежной работе алгоритма. Поэтому предлагается проведение преобразования изображения от яркостного к градиентному с последующей записью значений и направлений для каждого пикселя в соответствии с правилом

$$(Pix_{x_L, y_L}, Pix_{x_R, y_R}) \rightarrow (\|\nabla_{x_L, y_L}\|, \|\nabla_{x_R, y_R}\|),$$

где Pix_{x_L, y_L} — пиксели левого изображения;

Pix_{x_R, y_R} — пиксели правого изображения;

$\|\nabla_{x_L, y_L}\|$ и $\|\nabla_{x_R, y_R}\|$ — норма (длина) вектора градиента точки (x, y) на левом и на правом снимках.

На границах между темными и светлыми областями вектор градиента указывает направление перехода от темной к светлой области, а в его модуле содержится информация о фрагменте изображения. При этом значения градиентов вычисляются при помощи свёрток изображения с масками Собеля [3].

Так, для фрагмента изображения

$$\begin{bmatrix} f_{i-1, j-1} & f_{i-1, j} & f_{i-1, j+1} \\ f_{i, j-1} & f_{i, j} & f_{i, j+1} \\ f_{i+1, j-1} & f_{i+1, j} & f_{i+1, j+1} \end{bmatrix},$$

для элемента $f_{i, j}$ с масками

$$M_x = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}; \quad M_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

координаты вектора градиента Собеля

$\|\nabla\| = \sqrt{\nabla_x^2 + \nabla_y^2}$ имеют вид

$$\nabla_x = (f_{i-1, j+1} + 2f_{i-1, j} + f_{i-1, j-1}) - (f_{i+1, j-1} + 2f_{i+1, j} + f_{i+1, j+1})$$

$$\nabla_y = (f_{i-1, j-1} + 2f_{i, j-1} + f_{i+1, j-1}) - (f_{i-1, j+1} + 2f_{i, j+1} + f_{i+1, j+1})$$

Сумма всех весовых элементов масок Собеля равна нулю, а потому фон (т.е. участки с постоянной яркостью) дает нулевой отклик фильтра.

Чтобы найти некоторый пиксел (ячейку) p левого изображения с координатами (x_p, y_p) и нормой вектора градиента $\|\nabla_p\|$ на правом изображении, вся область перекрытия правого изображения должна быть проверяется на наличие пиксела p .

Предположим, что каждое значение градиента присутствует на изображениях с одной и той же вероятностью. Для устранения неоднозначности отождествления используются эталоны с размерами большими, чем один пиксел. Чем больше размеры эталона, тем сильнее он отличается от других участков изображения длинами векторов градиентов; тем самым обеспечивается однозначность. Чтобы повысить вероятность правильного отожде-

ствления, выбираются наиболее информативные эталоны с характерными особенностями изображения.

После этого поиск характерных точек производится методом корреляционного стереотождествления [4].

При этом каждая точка прямоугольной области

$$R' = \{(x_R, y_R) : x_{\min} \leq x_R \leq x_{\max}; y_{\min} \leq y_R \leq y_{\max}\}$$

на правом изображении поиска берется в качестве центра изображения, которое сравнивается с эталоном.

Мера сходства определяется корреляционной функцией [1]

$$k(x_R, y_R) = \frac{1}{\left(\sum_{(x,y)} v^2(x+x_L, y+y_L) - Nv^2 \right)^{1/2}} \times \frac{\sum_{(x,y)} v'(x+x_L, y+y_L)g'(x+x_R, y+y_R) - Nv'g'}{\left(\sum_{(x,y)} g^2(x+x_R, y+y_R) - Ng'^2 \right)^{1/2}},$$

где (x_L, y_L) — координаты центра эталона, выбранного на левом изображении;

(x_R, y_R) — координаты центра образа, выбранного на правом изображении;

$v'(x+x_L, y+y_L)$ — функция длины вектора градиента на эталонном участке изображения;

$g'(x+x_L, y+y_L)$ — функция длины вектора градиента изображения;

$\sum_{(x,y)}$ обозначает суммирование по всем ячейкам(пикселям) эталона;

N — количество ячеек в эталоне.

При этом средняя длина векторов градиентов будет находится по формулам

$$v' = \frac{1}{N} \sum_{(x,y)} v(x+x_L, y+y_L);$$

$$g' = \frac{1}{N} \sum_{(x,y)} g(x+x_R, y+y_R).$$

Тогда задача стереотождествления формулируется следующим образом: найти значения координат (x_R, y_R) , при которых корреляционная функция достигает максимума [6]:

$$(x_R^*, y_R^*) = \arg \max_{x_R, y_R} k(x_R, y_R). \quad (1)$$

Чем более информативным является эталон, тем сильнее выражен глобальный максимум по сравнению с локальными максимумами, а следовательно, точнее определяется область поиска R' , которая является прямоугольной областью поиска на правом изображении [5].

Один из способов исключения аномальных ошибок перепутывания локальных и глобального максимума состоит в увеличении размеров эталона, что в свою очередь увеличивает время стереотождествления и уменьшает его влияние на точность. Поэтому предлагается использовать пирамиду стереопар для ограничения области поиска R' и начальное приближение координат максимума корреляционной функции для исходного изображения.

Введение пирамидальной структуры данных преследует две основные цели:

- сокращение времени обработки изображений;
- определение более точных начальных приближений для обработки нижних уровней по результатам обработки верхних уровней.

Изображение сцены может быть представлено в различных пространственных масштабах. При этом крупные детали сцены лучше видны на изображениях с мелким (грубым) разрешением. Мелкие же детали сцены проявляются только на изображениях с высоким разрешением. Информативность участков изображения также зависит от разрешения. Уменьшение масштаба изображения по оси X (использование изображения с более грубым разрешением) приводит к увеличению скорости возрастания дисперсии при неизменных размерах области поиска, что позволяет использовать изображения с грубым разрешением для выбора эталонов, обладающих большой информативностью, повышающей качество стереотождествления.

Принцип построения пирамиды изображений показан на рис. 1.

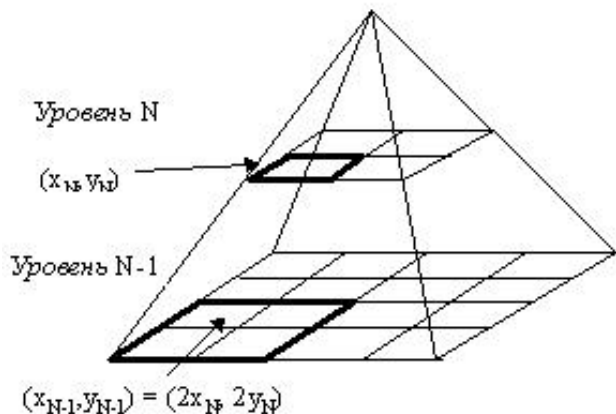


Рис. 1. Принцип построения пирамиды изображений

Каждое последующее изображение получается из предыдущего путем фильтрации и прореживания по каждой из осей в два раза (рис. 2).

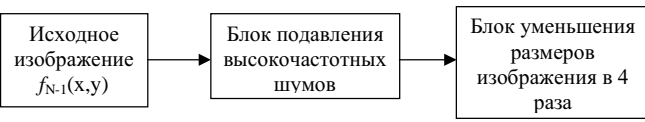


Рис. 2. Схема построения пирамиды изображения

Блок подавления высокочастотных шумов осуществляет фильтрацию в соответствии с правилом

$$b(x, y) = \iint h(u, v) f_{N-1}(x-u, y-v) du dv,$$

где $h(u, v)$ — ядро преобразования, являющееся функцией Гаусса, выбор которой обусловлен теоремой Котельникова, что обеспечивает сжатие с минимальной потерей информации.

На рис. 3 приведен результат использования пирамиды изображений, состоящей из пяти уровней

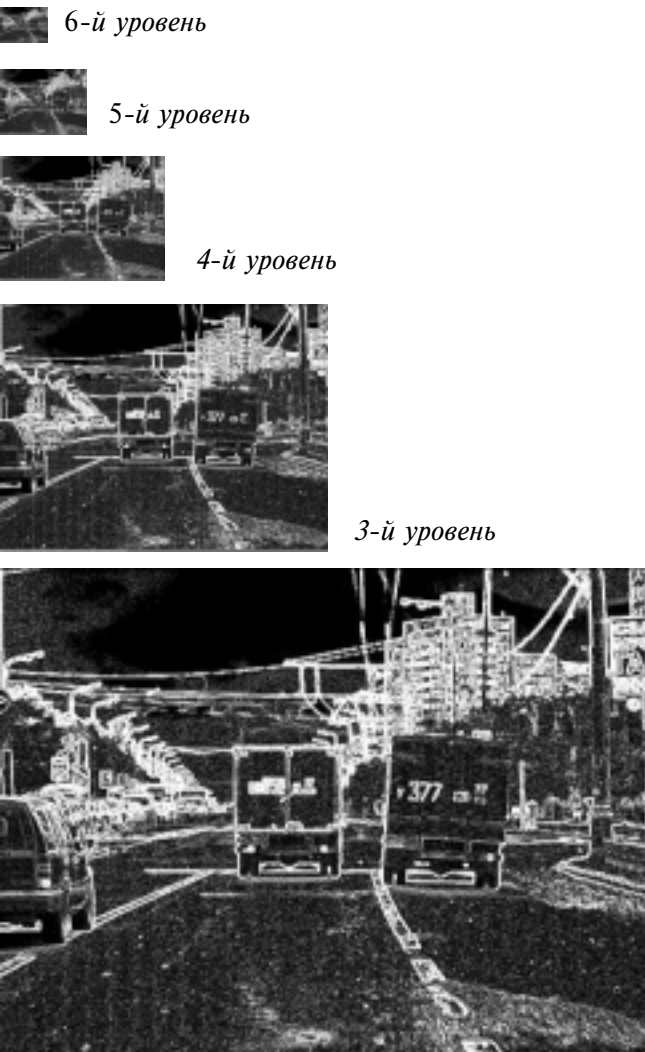


Рис. 3. Результат использования пирамиды из пяти уровней для тестового изображения после свертки маской Собеля

ней (исключая исходный — первый) для правого изображения тестовой стереопары.

Изображение N -го уровня представляет собой уменьшенную копию исходного изображения. Например, для уровня N размер пикселя изображения вычисляется следующим образом: $p_n = 2^{N-1}$, а для координат пикселей изображений двух произвольных уровней пирамиды с номерами n и m справедливы следующие соотношения:

$$2^{n-1} x_n = 2^{m-1} x_m, \quad 2^{n-1} y_n = 2^{m-1} y_m.$$

Тогда при использовании пирамиды изображений на N -м уровне площадь области поиска R' сокращается в $(2^{N-1})^2$ раз по сравнению с площадью области поиска на 1-м уровне:

$$R'_N = \left\{ (x_N, y_N) : \frac{x_{\min}}{2^{N-1}} \leq x_N \leq \frac{x_{\max}}{2^{N-1}}, \frac{y_{\min}}{2^{N-1}} \leq y_N \leq \frac{y_{\max}}{2^{N-1}} \right\}.$$

Пусть на N -м уровне осуществлено стереотождество в точке $(x^*, y^*)_N \in R'_N$, тогда эту точку можно рассматривать в качестве центра области поиска на $(N-1)$ -м уровне.

Время работы алгоритма стереотождество пропорционально количеству ячеек (пикселей) в области поиска. Без применения пирамиды изображений время работы алгоритма определяется соотношением

$$T_1 = \alpha S(R') = \alpha (x_{\max} - x_{\min} + 1)(y_{\max} - y_{\min} + 1),$$

где $S(R')$ — количество ячеек (пикселей) внутри области R' , а при использовании пирамиды изображений — соотношением

$$T_2 = \alpha \sum_{i=1}^N S(R'_i) = \alpha \frac{(x_{\max} - x_{\min} + 1)(y_{\max} - y_{\min} + 1)}{(2^{N-1})^2} + \alpha(N-1)(2\Delta_x + 1)(2\Delta_y + 1),$$

2-й уровень

где Δ_x, Δ_y — константы, характеризующие размер области поиска.

Анализ этих соотношений показывает, что чем больше область поиска и чем больше уровней содержится в пирамиде, тем быстрее работает алгоритм стереотождествления [4]. При значениях констант $x_{\max} - x_{\min} + 1 = 100, y_{\max} - y_{\min} + 1 = 5, N = 4, \Delta_x = 2, \Delta_y = 1$ применение пирамиды стереопар дает выигрыш во времени $T_1 / T_2 \approx 5$.

Для каждого эталона, выбранного на левом изображении, на всех уровнях пирамиды решается задача стереотождествления (1). Область R'_N и область R'_i ($i = N - 1, \dots, 1$) определяются выражением с использованием начального приближения $(x_R^*, y_R^*)_i$. На самом нижнем уровне (который представляют собой исходные изображения) получаются значения $(x_R^*, y_R^*)_1$, которые определяют положение образа, соответствующего эталону [2]. Задача стереотождествления решается последовательным перебором значений параметров сдвига из допустимого диапазона. На более детальных уровнях используется начальное приближение, полученное с предыдущих уровней.

Зависимость времени T (мс) на проведение стереотождествления точки от количества уровней N в пирамиде изображений показана на рис. 4.

Результатом работы алгоритма стереотождествления (рис. 5) является надежное определение точки левого снимка, найденной на исходном изображении за время, затрачиваемое на обработку одной стереопары и равное 18 мс. Операция проводилась с использованием тестового автомобиля-лаборатории

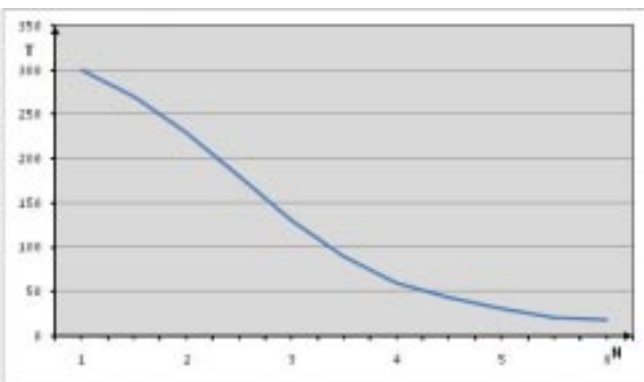


Рис. 4. График зависимости времени на стереотождествление от количества уровней пирамиды



Рис. 5. Пример определения точки (x_1, y_1) левого снимка на правом (x_2, y_2)

с расположенными в нем под лобовым стеклом двумя ПЗС-камерами, образующими стереопару.

Таким образом, использование пирамиды изображений позволяет ускорить процесс стереотождествления, а использование карты градиентов вместо информации о яркости способствует надежному обнаружению точки на подстилающей поверхности, не имеющей характерных особенностей рельефа.

Summary

Full time needed for stereo identification is intolerably large in problems of stereopair-based video measurement aimed to motion control for dynamic systems. An algorithm is suggested to overcome this situation. The algorithm uses some gradient approach and an image pyramid to perform stereo identification for pixels obtained from two cameras with mutually known orientation parameters. The algorithm is implemented on experimental facility equipped with

CCD cameras. The resolution is 640x480 pixels and the frame rate is 30 frames per second for the cameras.

Библиографический список

1. Желтов С.Ю., Степанов А.А., Визильтер Ю.В., Морзеев Ю.В. Объектно-ориентированный фрейм-овый подход в обработке изображений и управления данными, получаемыми от датчиков различной физической природы // SPEE Proceedings. 1999. V. 2597. № 4. P. 302-303.

2. Тарасов В.Б., Степанов А.А., Желтов С.Ю. Нечеткие модели в обработке изображений: обзор зарубежных достижений // Новости искусственного интеллекта. 1993. № 3. С. 40-64.

3. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. — М.: ИД «Вильямс», 2004.

4. Себряков Г.Г., Визильтер Ю.В. Разработка методики построения специализированных экспертных систем для анализа цифровых изображений в задачах обнаружения и идентификации сложных структурных объектов // Вестник компьютерных и информационных технологий. 1997. № 3. С. 31-32.

5. Красовский А.А., Белоглазов И.Н., Чигин Г.П. Теория корреляционно-экстремальных навигационных систем. — М.: Наука, 1979. — С. 448-452.

6. Чеголин П.М., Леонович Э.Н., Савенков В.И. Автоматизация преобразования сложных форм графической информации. — Минск: Наука и техника, 1973.

Московский авиационный институт,
ФГУП «ГосНИИАС»

Статья поступила в редакцию 30.03.2008