

Автономная навигация в закрытых помещениях с использованием адаптивного описания эталонных изображений

Л.И. Лебедев, Ю.Г. Васин, М.П. Осипов¹

lebedev@pmk.unn.ru|pmk@unn.ac.ru|osipovmp@mail.ru

¹ Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия

В работе описываются методы коррекции результатов позиционирования посредством автоматического распознавания фрагментов графических изображений, инвариантного относительно аффинных преобразований. Распознавание осуществляется на базе корреляционно-экстремального контурного метода. В целях увеличения быстродействия и качества распознавания предлагается метод адаптивного формирования описаний эталонных изображений в зависимости от параметров проецирования. Приводится оценка сложности предлагаемого метода распознавания.

Ключевые слова: распознавание, коррекция, аффинное преобразование, эталонные изображения, адаптивное описание, видео поток, КЭКМ.

Autonomous navigation in enclosed spaces using an adaptive description of standard images

L.I. Lebedev, Yu.G. Vasin, M.P. Osipov¹

lebedev@pmk.unn.ru|pmk@unn.ac.ru|osipovmp@mail.ru

¹ N.I. Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia

The paper describes methods for correcting the results of positioning by automatically recognizing fragments of graphic images that are invariant with respect to affine transformations. Recognition is carried out on the basis of the correlation-extreme contour method. In order to increase the speed and quality of recognition, a method is proposed for the adaptive formation of descriptions of standard images, depending on the projection parameters. The complexity of the proposed recognition method is estimated.

Keywords: recognition, correction, affine transformation, standard image, adaptive description, video stream, CECM.

1. Введение

В настоящее время актуальной темой исследовательских работ является поиск доступных технологических решений в задаче навигации внутри зданий и помещений.

Классический подход к навигации с использованием технологий спутниковых радионавигационных систем на основе GPS и ГЛОНАСС малоэффективен внутри помещений поскольку, стены и перекрытия выступают помехой для прохождения спутникового сигнала. Поэтому на данный момент предлагаются различные подходы решения задачи навигации в закрытых помещениях, использующие: сети передачи данных (Bluetooth, Wi-Fi, FM-радиоволны, UWB), магнитное поле Земли, микромеханические инерционные датчики (акселерометры, датчики угловой скорости), QR-коды, RFID-метки и т.д. [7]. Но каждый подход в отдельности имеет существенные ограничения по точности, надежности, стоимости и сложности реализации.

Традиционным способом навигации внутри помещений на данный момент остается визуальная навигация, представляющая собой систему указателей и именных табличек, размещенных в здании и служащих путеводителем по нему. Несмотря на простоту реализации, существенный недостаток такого подхода состоит в большой информационной нагрузке на человека и сложности анализа полученной информации. В последнее время в ряде работ было предложено использовать методы компьютерного зрения как для анализа движения [4], так и для анализа информационных объектов помещения (указателей, именных табличек и т.д.) [9], что позволяет автоматизировать процесс навигации. На стенах и дверях зданий часто можно встретить плоские объекты,

указывающего и предупреждающего характера. Это названия и номера помещений, знаки пожарной безопасности, путей эвакуации, указатели и т.д. Такие информационные объекты имеют яркую окантовку, что легко выделяет их из общего фона. Если, используя методы компьютерного зрения, распознать такие плоские объекты в кадре с камеры мобильной навигационной системы, то зная точное расположение этих объектов можно восстановить информацию о местоположении пользователя. Для этого предварительно создается база данных изображений плоских информационных объектов помещения (картины, названия и номера помещений, планы эвакуации, знаки пожарной безопасности и т.д.) с координатной привязкой их расположения. Таким образом, в момент, когда необходима корректировка своего местоположения, пользователь производит видеосъемку близлежащего плоского информационного объекта с помощью мобильного устройства со встроенной камерой. Далее, используя методы поиска изображения по содержанию (англ. Content-based image retrieval (CBIR)), происходит сравнение элементов изображения полученных при видеосъемке с базой данных изображений информационных объектов. Если искомое изображение найдено, то, зная информацию о расположении изображаемого объекта, хранящуюся в базе данных, можно определить расположение и ориентацию камеры, производящей съемку, а, следовательно, расположение и ориентацию пользователя. Недостатком такого подхода является не уникальность информационных объектов, а следовательно неоднозначность определения местоположения пользователя.

В предыдущей работе авторов [3] для решения задачи навигации внутри здания был применен комбинированный подход, в основе которого лежат методы инерциальной

навигации [5] и компьютерного зрения [12] при использовании поэтажных планов здания, технологий моделирования и 3D визуализации. Использование комбинированного подхода позволяет компенсировать недостатки каждого метода и повысить надежность позиционирования. Для распознавания информационных объектов помещения использовался метод SIFT поиска локальных особенностей для определения соответствий на изображении [6]. Этот метод традиционно применяется исследователями при использовании методов компьютерного зрения в задаче навигации [9,12]. Но метод SIFT (SURF) имеет высокую вычислительную сложность и плохо работает для объектов простой формы и без ярко выраженной текстуры какими, как правило, являются информационные объекты помещения. В работе [8] для распознавания текстовой информации в табличках на стенах помещения в задаче навигации используется метод Optical Character Recognition (OCR). Но данный метод не является инвариантным относительно ортогональных преобразований и масштабирования, что значительно сокращает возможности его применения в задаче навигации. В свою очередь, информационные объекты помещения имеют четкие границы и легко выделяются из окружающего фона, поэтому поиск такого класса объектов можно эффективно производить методами контурного анализа. В качестве метода контурного анализа предлагается использовать корреляционно-экстремальный контурный метод (КЭКМ) [2].

2. Постановка задачи

На базе эталонных изображений уникальных информационных объектов должна решаться задача автоматического распознавания элементов текущего кадра из видеопотока с последующей его идентификацией. Под уникальностью информационных объектов здесь понимается как содержащееся в них информация, так и возможность ее распознавания предлагаемыми методами. Кроме того, для обеспечения быстродействия и качества распознавания, предлагаемые методы должны быть инвариантными относительно аффинных преобразований, а также к адаптивному описанию фрагментов изображения в зависимости от угла сканирования информационного объекта. Коррекция результатов позиционирования должна осуществляться на основе данных о местоположении эталонного изображения и параметров сканирования, полученных в результате распознавания информационного объекта.

3. Методы решения

Распознавание изображений информационных объектов будет осуществляться методами сравнения с эталонами посредством обнаружения и идентификации фрагментов, сходных по описанию с элементами эталонных изображений. Эталонное изображение формируется на основе описаний фрагментов изображения плоского информационного объекта, являющегося центральной фронтальной перспективой. Совокупность элементов эталонного изображения должна быть достаточной для качественного автоматического распознавания информационных объектов с одной стороны и минимальной по составу для обеспечения необходимого быстродействия с другой стороны.

Распознавание фрагментов изображения будем осуществлять корреляционно-экстремальным контурным методом (КЭКМ) [2]. В основу КЭКМ положено вычисление оценки близости ε_m по описаниям $\mathbf{w} = (x, y)^T$ фрагмента изображения $\mathbf{O}$ и $\mathbf{w}^e = (x^e, y^e)^T$

эталона $\mathbf{E}$, которые задаются посредством векторной (контурной) модели представления информации:

$$\varepsilon_m = \mathbf{Dw}^e - R^2 / \mathbf{Dw}, \quad (1)$$

где $\mathbf{Dw}^e$, $\mathbf{Dw}$ - дисперсии контуров эталона и фрагмента изображения соответственно, а R - величина, вычисленная по значениям смешанных корреляционных моментов:

$$\begin{aligned} R^2 &= S_n^2 + C_s^2, \\ S_n &= \text{cov}(\mathbf{x}^e, \mathbf{y}) - \text{cov}(\mathbf{y}^e, \mathbf{x}), \\ C_s &= \text{cov}(\mathbf{x}^e, \mathbf{x}) + \text{cov}(\mathbf{y}^e, \mathbf{y}). \end{aligned} \quad (2)$$

Начальные моменты первого и второго порядка находятся по формулам:

$$\begin{aligned} \bar{\mathbf{w}}^e &= \mathbf{Mw}^e = \frac{1}{2 \cdot S^e} \cdot \sum_{i=1}^{m-1} (\mathbf{w}_i^e + \mathbf{w}_{i+1}^e) \cdot S_i^e \\ \mathbf{My}^2 &= \frac{1}{3 \cdot S} \cdot \sum_{i=1}^{n-1} (y_i^2 + y_{i+1}^2 + y_i \cdot y_{i+1}) \cdot S_i \end{aligned} \quad (3)$$

где $\mathbf{w}^e = \{\mathbf{w}_1^e, \mathbf{w}_2^e, \dots, \mathbf{w}_m^e\}$, $\mathbf{w} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$, $\mathbf{w} = (x, y)^T$ - исходные описания контуров эталона и распознаваемого фрагмента соответственно, а S_i^e , S_i - длины ребер контуров. Вычисление смешанных моментов второго порядка осуществляются по вспомогательным описаниям эталона $\hat{\mathbf{w}}_j^e = (\hat{x}_j^e, \hat{y}_j^e)^T$ и фрагмента $\hat{\mathbf{w}}_j = (\hat{x}_j, \hat{y}_j)^T$, $j = 1, 2, \dots, l \leq m + n - 2$ соответственно [8]:

$$\begin{aligned} \mathbf{Mx}^e \mathbf{y} &= \frac{1}{6 \cdot S^e} \cdot \sum_{i=1}^{l-1} (2 \cdot \hat{x}_i^e \cdot \hat{y}_i + 2 \cdot \hat{x}_{i+1}^e \cdot \hat{y}_{i+1} + \\ &\quad + \hat{x}_i^e \cdot \hat{y}_{i+1} + \hat{x}_{i+1}^e \cdot \hat{y}_i) \cdot \hat{S}_i^e \end{aligned} \quad (4)$$

Оценка близости ε_m корректна только для согласованных описаний контуров. Для согласования описаний контуров применяются эффективные по быстродействию методы относительных смещений и метод парабол [2, 11]. Поэтому, КЭКМ инвариантен не только относительно ортогональных преобразований и масштабирования, но также и к циклическим описаниям контуров, что обеспечивает наиболее полный состав элементов в классах эквивалентностей.

В работе [10] было доказано, что любое аффинное преобразование может быть представлено совокупностью операций, включающих проецирование контура на плоскость, полученную последовательным вращением плоскости XOY вокруг осей OX и OY на углы β и γ соответственно, ортогональное преобразование и масштабирование проецированного контура. Для обеспечения инвариантности относительно аффинных преобразований вначале формируются описания проекций контура по формулам

$$\begin{cases} x_{pr} = x \cdot \cos \gamma + y \cdot \sin \beta \cdot \sin \gamma \\ y_{pr} = y \cdot \cos \beta \end{cases}, \quad (5)$$

а затем решается оптимизационная задача вида

$$\varepsilon_m^A = \min_{(\beta, \gamma) \in D^o} \varepsilon_m(\beta, \gamma), \quad (6)$$

где оценка сходства (близости) $\varepsilon_m(\beta, \gamma)$ для заданных значений углов β и γ находится с помощью КЭКМ в соответствии с формулами (1)-(4). Решение

оптимизационной задачи (6) осуществляется с использованием метода сеток. Область $D^\circ(\beta, \gamma)$ изменения параметров β и γ может быть изначально детализирована при распознавании контуров, описывающих специфические фрагменты изображений (рамки, условные знаки и т.д.) при непосредственном вычислении матрицы аффинного преобразования A° . Область D° в процессе распознавания фрагментов информационных объектов может корректироваться. Необходимо также отметить, что нахождение области D° позволяет значительно увеличить быстродействие алгоритма автоматического распознавания.

4. Полученные результаты

В результате исследований различных структур описания эталонных изображений была выбрана модель представления в формате КМР, адаптированная для решения задачи распознавания информационных объектов. Предлагаемая структура описания эталонного изображения включает три поля: заголовок, параметрическое описание каждого из эталонов и их метрические описания путем задания последовательностей координат точек контуров. Заголовочная часть содержит сведения об эталонном изображении в целом: идентификатор эталонного изображения, количество эталонов N_1 , представляющих изображение, координаты привязки изображения (x, y, H) , разрешение и ряд других сведений об изображении. Атрибутивное описание эталонов состоит из записей постоянной длины и включает данные о количестве точек в метрическом представлении эталона, его габаритах, местоположении эталона на изображении информационного объекта, ссылки на адрес начала метрического описания эталона, типа и кода эталона, раскраска контура. Также, вычисляются и запоминаются дисперсионные характеристики эталона и ряд других сведений, необходимых для обеспечения быстрого действия при нахождении согласованных описаний. Кроме того, для каждого контура эталонного изображения задаются условия его активизации и значимости. Это позволяет в зависимости от условий сканирования информационного объекта изменять необходимый набор эталонов, обязательных для опознавания полученного изображения.

Для получения контурной модели представления изображения были реализованы несколько алгоритмов, которые последовательно обеспечивают перевод его в оттенки «серого», затем, осуществляют бинаризацию с порогом Th , полученным с использованием метода Оцу, далее решают задачу получения штриховой модели изображения и, наконец, построения самой контурной модели в соответствии с методом формирования КМР [1]. В совокупности все представленное алгоритмическое и программное обеспечение построения контурной модели обладает необходимым для решения задачи распознавания изображений информационных объектов быстродействием.

Формирование эталонного изображения осуществляется на основе полученной контурной модели базового информационного объекта в интерактивном режиме. Пользователю предоставляется возможность из множества контуров, описывающих изображение в формате КМР выбрать необходимую из них совокупность, достаточную для идентификации изображения этого информационного объекта, получаемого в других условиях сканирования. Уменьшение угла сканирования информационного объекта влечет изменение состава контуров в описании изображения в формате КМР в

результате их слипания. Поэтому, при формировании эталонного изображения предусмотрена возможность получения описаний слипшихся контуров в базовых изображениях, чтобы использовать их в качестве эталонов, что способствует повышению качества распознавания. Параметры, характеризующие ранговые предпочтения контуров и их условия использования в обязательном порядке задаются пользователем также интерактивным режиме. Остальные атрибуты описания контура формируются автоматически.

Для распознавания информационных объектов был реализован алгоритм на базе КЭМ. Информационный объект будет считаться распознанным, если на нем будут идентифицированы все контура, соответствующие элементам эталонного изображения с учетом их представления. Особенностью данного алгоритма распознавания является то, что он может использовать как прямое, так и обратное аффинное преобразование. Прямое аффинное преобразование, обеспечивающее переход от базового описания изображения к текущему, необходимо для локализации области поиска контуров, соответствующих местоположению рассматриваемого эталона. Это исключает необходимость распознавания всех контуров на изображении текущего информационного объекта. Обратное аффинное преобразование дает возможность при вычислении оценки сходства использовать методы относительных смещений и парабол для получения согласованных описаний контуров. Это также направлено на увеличение быстродействия алгоритма распознавания.

На примере распознавания конкретного информационного объекта, кадр с изображением которого представлен на рис. 1, представим все этапы, необходимые для идентификации и проведения коррекции результатов позиционирования. Отметим, что это полученное с камеры изображение является боковой фронтальной перспективой.



Рис. 1. Кадр с изображением информационного объекта.

Вначале для данного информационного объекта сформируем эталонное изображение. Очевидно, что заранее невозможно определить угол, под которым будет получено распознаваемое изображение информационного объекта. Поэтому формирование эталонов будет осуществляться на основе элементов изображения, представленного на рис.2 и являющегося центральной фронтальной перспективой того же объекта. Изображение данного базового информационного объекта в формате контурной модели представлено на рис. 3. Представленная контурная модель изображения получена при бинаризации с порогом $Th = 120$. Как было отмечено выше, при формировании эталонного изображения необходимо использовать такие фрагменты базового изображения, чтобы они однозначно определяли данный информационный объект. Эксклюзивными элементами

данного информационного объекта являются символы, идентифицирующие ответственного за противопожарную безопасность и текст, характеризующий предметную область распознаваемого объекта. Поэтому, эталонное изображение должно включать, во-первых, описания рукописных букв 2-х «Л» (2 контура), 3-х «е» (2), «б» (1), «в» (1), «И» (1), а во-вторых, символы из фрагментов текста, например, «помещения», «пожаре» и «или». Метрические описания одних и тех же символов представляются одним контуром и при распознавании КЭМ заимствуются, что позволяет увеличить быстродействие алгоритма распознавания отдельных букв. При вращении базового изображения относительно оси OY на угол более 65° происходит слипание символов «д» и «е» в фамилии ответственного, а при угле более 77° - слипание символов «Л» и «е», «б» и «е». При превышении этих углов в первом случае слипания буква «е» перестает быть значимым эталоном, а во втором, эталонное изображение для более качественного распознавания можно пополнить вновь образовавшимися контурами. В результате получим эталонное изображение с адаптивным для распознавания составом эталонов.

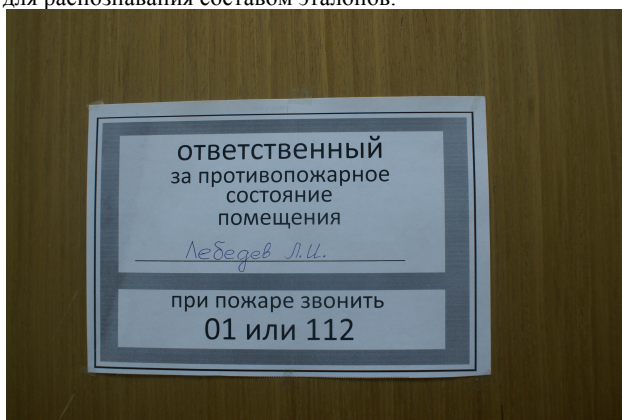


Рис. 2. Базовое изображение информационного объекта.

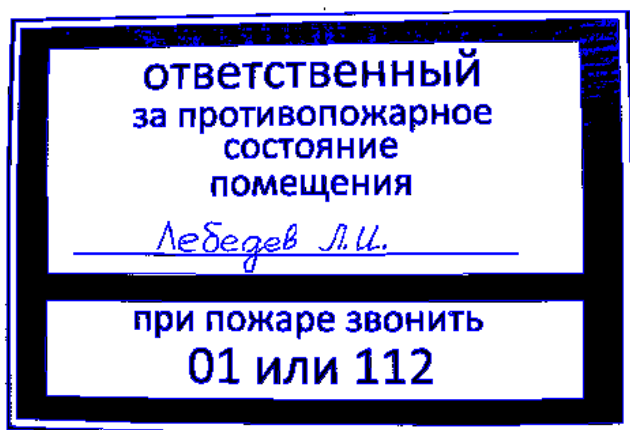


Рис. 3. Базовый информационный объект в формате КМР.

Изображение информационного объекта в формате КМР включает описание двух внутренних контуров, задающих границы рамок. Включение любого из этих контуров в описание эталонного изображения позволило легко получить матрицу аффинного преобразования

$$A^\circ = \begin{pmatrix} 0.3396 & -0.0776 \\ 0.0542 & 0.9846 \end{pmatrix},$$

по ее компонентам найти углы проецирования $\beta \approx 4.9^\circ$ и $\gamma \approx 69.6^\circ$, определить границы области $D^\circ(\beta, \gamma)$ и целенаправленно отбирать элементы информационного

объекта для распознавания, используя относительные координаты положения эталонов на базовом изображении.

На рис. 4 приведен фрагмент изображения текущего информационного объекта. Все символы данного изображения, соответствующие эталонам базового информационного объекта, были распознаны с вероятностью, значительно превосходящей пороговое значение. Исключение составляет третья буква «е» в фамилии, однако ее распознавание не осуществлялось алгоритмом из-за отсутствия соответствующего эталона согласно его области использования.

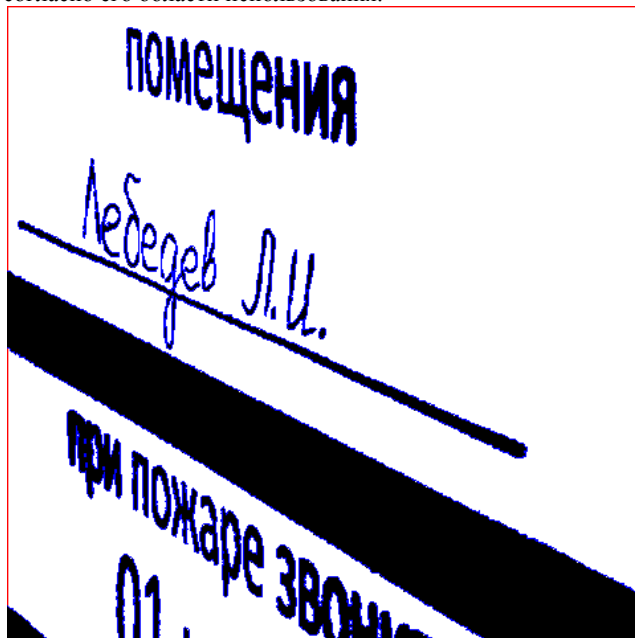


Рис. 4. Фрагмент КМР- изображения текущего кадра.

Таким образом, данное изображение было успешно распознано, а координаты его привязки использованы для коррекции результатов позиционирования.

Оценим вычислительную сложность предложенного подхода коррекции текущих координат позиционирования, которая полностью определяется сложностью распознавания изображения информационного объекта. Наиболее сложным в вычислительном плане является получение начальной матрицы аффинного преобразования

A° . Для ее получения в общем случае необходимо вычислить оценки сходства эталона со всеми N_2 контурами на изображении информационного объекта.

Таким образом, сложность вычисления матрицы A° оценивается величиной $\Xi_1 = \lambda \cdot N_2 \cdot O(m + \bar{n})$, $\lambda < 100$, где $\bar{n}$ - среднее количество точек в описаниях используемых контуров изображения информационного объекта. Отметим, что за счет логической фильтрации количество контуров N_2 можно значительно уменьшить.

После получения матрицы аффинного преобразования A° вычисление оценок сходства осуществляется только для эталонов и контуров на изображении информационного объекта, соответствующих им по местоположению. Сложность вычисления этих оценок сходств с учетом количества уточнений ϑ матрицы аффинного преобразования оценивается величиной $\Xi_2 = N_1 \cdot \vartheta \cdot O(\bar{m} + \bar{n})$, $(\vartheta \leq 6)$. Таким образом, сложность распознавания оценивается величиной $\Xi = \Xi_1 + \Xi_2$.

5. Заключение

Предлагаемый подход коррекции результатов позиционирования, ориентированный на использование алгоритма распознавания фрагментов графических изображений, реализованного на базе КЭМ и инвариантного относительно аффинных преобразований, увеличил возможности осуществления автономной навигации в закрытых помещениях. Дальнейшего увеличения качества автоматического распознавания можно достичь путем уменьшения перспективных искажений изображения.

6. Благодарности

Работа выполнена при поддержке Российского Фонда Фундаментальных исследований, проект № 16-07-01214 А и проект № 16-07-01198.

7. Литература

- [1] Васин Ю.Г., Башкиров О.А., Чудинович Б.М., Гервагин Ю.Г. Подсистема растрового ввода-вывода и обработки графической информации РАСТР I // Автоматизация обработки сложной графической информации. Межвузовский сборник / Горьковский государственный университет, г. Горький, 1984.
- [2] Лебедев Л.И. Корреляционно – экстремальные контурные методы распознавания. Теоретические основы: Учебное пособие./ Нижний Новгород, Изд-во Нижегородского государственного университета, 2013. 113 с. – ISBN 978-5-91326-308-7.
- [3] М.П. Осипов, А.О. Патрушев Методы корректировки местоположения в задаче автономной навигации в закрытых помещениях / Труды 26-ой Международной конференции по компьютерной графике и зрению, Нижний Новгород. 2016, С. 417-419.
- [4] S. Hilsenbeck, A. Moller, R. Huitl, G. Schroth, M. Kranz, and E. Steinbach, "Scale-preserving long-term visual odometry for indoor navigation," in International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, Australia, Nov. 2012.
- [5] Jeong Won Kim, A Step, Stride and Heading Determination for the Pedestrian Navigation System / Jeong Won Kim, Han Jin Jang, Dong-Hwan Hwang, Chansik Park // In: Positioning, Vol.1, No.8, December 2004 – pp.273-279.
- [6] Lowe, David G. (1999). "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. pp. 1150–1157.
- [7] Rainer Mautz, Indoor Positioning Technologies / Rainer Mautz -Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, February 2012 – c.127.
- [8] Sami M., Ayaz Y., Jamil M., Gillani S. O., Muhammad N. Text detection and recognition for semantic mapping in indoor navigation. International Conference on IT Convergence and Security. IEEE, 2015, pp. 1–4.
- [9] Serrão, M., Shahrabadi, S., Moreno, M., Josre, J. T., Rodrigues, J. I. Rodrigues, J. M. F., du Buf, J. M. H., Computer vision and GIS for the navigation of blind persons in buildings. Universal Access in the Information Society, 14(1), pp. 67-80, 2014.
- [10] Vasin Ju.G., Lebedev L.I. Recognition of base of similarity estimations, invariant with respect to affinities. //8th International conference on Pattern recognition and image analysis: new information technologies"/ Yoshkar-Ola, v.2, 2007,.pp. 57-60.
- [11] Vasin Yu.G., Lebedev L.I. The problem of obtaining coherent contour descriptions in the calculation of similarity estimates. //8th Open German-Russian Workshop "Pattern recognition and image understanding" (OGRW-8-2011): Workshop Proceedings. 2011. P. 324-327.
- [12] Werner M., Kessel M., and Marouane C., Indoor positioning using smartphone camera, in Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '11), 6, p. 1, September 2011.

Об авторах

Лебедев Леонид Иванович, к.ф.-м.н., ст.н.с., ведущий научный сотрудник института информационных технологий, математики и механики Нижегородского государственного университета им. Н.И. Лобачевского. Его e-mail: lebedev@pmk.unn.ru.

Васин Юрий Григорьевич, д.т.н., главный научный сотрудник института информационных технологий, математики и механики Нижегородского государственного университета им. Н.И. Лобачевского. Его e-mail: pmk@unn.ac.ru.

Осипов Михаил Павлович, к.т.н., ведущий научный сотрудник института информационных технологий, математики и механики Нижегородского государственного университета им. Н.И. Лобачевского. Его e-mail: osipovmp@mail.ru.