

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФОРМИРОВАНИЯ ОБРАЗОВ

УДК 004.021

ОПТИМАЛЬНАЯ АФФИННАЯ АППРОКСИМАЦИЯ ПРОЕКТИВНОГО ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЙ

© 2019 г. И. А. Коноваленко^{1,2,*}, В. В. Кохан¹, Д. П. Николаев^{1,2}

¹ Институт проблем передачи информации им. А.А. Харкевича РАН
127051 Москва, Большой Каретный пер., 19, Россия

² Smart Engines Ltd., 117312 Москва, проспект 60-летия Октября, д. 9, Россия

*E-mail: konovalenko@iitp.ru

Поступила в редакцию 06.09.2018 г.

Замена проективного преобразования существенно более простым аффинным преобразованием встречается во многих областях технического зрения. В то же время понятие точности аффинной аппроксимации проективного преобразования в литературе не формализуется. Это в свою очередь приводит к отсутствию постановок задач и теоретически обоснованных методов аффинной аппроксимации проективного преобразования. Настоящая работа направлена на заполнение этого пробела. Авторами предложено в качестве критериев точности аффинной аппроксимации проективного преобразования изображения использовать среднеквадратичную и максимальную поточечную невязку в координатах преобразованного изображения. На основе этих критериев сформулированы задачи поиска оптимальных аффинных аппроксимаций. Доказана выпуклость полученных задач оптимизации. Предложен метод использования оптимальных аффинных аппроксимаций для экономии вычислительных ресурсов при преобразовании изображений.

Ключевые слова: проективное преобразование, точность проективного преобразования, аффинная аппроксимация преобразований, линеаризация, выпуклый анализ

DOI: 10.1134/S0235009219010062

ВВЕДЕНИЕ

Проективное исправление

Рассмотрим произвольную оптическую систему. Оптическая система обладает дисторсией, если коэффициент линейного увеличения изменяется по полю зрения ее объекта. Оптическая система, свободная от дисторсии, называется ортоскопической оптической системой (ООС). Любую оптическую систему можно с произвольной точностью свести к ООС посредством процедуры калибровки, поэтому далее будем рассматривать только ООС. Они обладают тем важнейшим свойством, что образы прямых линий сцены, которые они порождают, тоже являются прямыми линиями или точками.

Под коррекцией перспективы порожденного ООС изображения будем понимать такое геометрическое преобразование этого изображения, которое устраняет на нем перспективное схождение прямых линий, параллельных в сцене. Будем рассматривать частный случай, когда сцену можно считать плоской. Тогда коррекция перспективы существует и является проективным преобразованием. Поэтому коррекцию перспективы для случая плоской сцены будем называть проективным исправлением.

Большое количество прикладных и теоретических исследований на тему проективного исправления подтверждает ее актуальность. Например, в работах (Orrite, Herrero, 2004; Zhukovsky et al., 2015) предлагаются алгоритмы для проективного сопоставления контуров и отрезков соответственно. В работе (Povoltskiy et al., 2017) описывается сегментация проективно исправленных номеров знаков автомобилей, а в (Чернов и др., 2016) – паспортов. В статьях (Шемякина и др., 2017; Шемякина, 2017) анализируются подходы к решению задачи проективного наведения на примере страховых свидетельств. В работах (Болотова и др., 2017; Skoryukina et al., 2018) рассматриваются алгоритмы проективного наведения на документы. В статье (Skoryukina et al., 2017) описывается система “Snapscreen”, которая детектирует и проективно исправляет фотографию экрана телевизора с дальнейшим распознаванием телевизионной передачи. В статьях (Kargenka et al., 2015; Холопов, 2017) задача проективного исправления возникает при картировании местности по снимкам с бортовой камеры БПЛА. Теоретические исследования проективных инвариантов представлены, например, в работах (Николаев, 2016; Савчик, Николаев, 2016; Балицкий и др., 2017).

Аффинная аппроксимация проективного преобразования

Мы будем исследовать замену проективного преобразования существенно более простым аффинным преобразованием. Соображение, что на практике аффинное преобразование с хорошей точностью локально приближает проективное преобразование, было высказано в статье (Gruen, 1985). Это свойство было использовано в работе (Tu-ichi et al., 1981) для упрощения дальнейших математических построений. Аффинное приближение повсеместно используется в рендеринге изображений (Heckbert, 1989; Lorenz, Dollner, 2009) и в завершении изображений (Darko, 2006). В работе (Huang et al., 2015) проективное преобразование заменяется на более простое аффинное с целью борьбы с переобучением. Близкая идея используется в “слабой перспективной модели” (Alter, 1992; Kutulakos, Vallino, 1996; Aradhye, Myers, 2010), где аппроксимация носит частичный характер. Строить аффинно инвариантные методы (вместо существенно более сложных проективно инвариантных) принято в популярной технологии особых точек (Mikolajczyk, Schmid, 2002; Mikolajczyk, Schmid, 2004; Morel, Yu, 2009), как и в родственной задаче поиска значимых регионов (Kadir et al., 2004), хотя оба подхода практически обеспечивают инвариантность к ракурсу съемки. Деление на аффинные и проективные методы имеется и в области стереорекострукции (Faugeras, 1992).

Существует критика аффинной аппроксимации, как ведущей к практической потере точности (Zwicker, 2004). Однако само понятие точности аффинной аппроксимации проективного преобразования в литературе не формализуется. Это в свою очередь делает невозможным поиск оптимальной аффинной аппроксимации проективного преобразования. Основное содержание данной работы состоит в введении понятия точности аффинной аппроксимации проективного преобразования и обоснования вычислительной возможности нахождения оптимальной по точности аффинной аппроксимации.

Критерии точности проективного исправления

Пусть дано два проективных преобразования H и $\hat{H}$. Первое — имеет смысл идеального исправления входного изображения, которым обычно является фотография. На практике H неизвестно, но при тестировании может задаваться экспертно. Второе — некоторое практическое приближение или оценка первого, как правило, получаемое автоматически. В результате применения H к входному изображению получается некоторое изображение, называемое идеально исправлен-

ным, а в результате применения $\hat{H}$ — практически исправленным.

Кроме того, пусть дана область интереса R — множество тех и только тех точек идеально исправленного изображения, в которых нам важно, чтобы $\hat{H}$ хорошо приближало H . Обычно R соответствует документу, автомобильному номеру, зданию и прочим интересующим нас объектам сцены, либо же задает идеально исправленное изображение целиком. Множество R известно, постоянно для конкретной задачи и не зависит от H и $\hat{H}$. Например, при проективном исправлении фотографии водительских прав мы заранее задаем, какая область изображения при идеальном его исправлении должна оказаться занятой образом прав.

Также введем зависимое множество $Q \stackrel{\text{def}}{=} \hat{H}H^{-1}(R)$, содержащее образы тех же точек сцены, что и R , но на практически исправленном изображении. Тогда встает задача формализации точности приближения преобразованием $\hat{H}$ преобразования H . Наиболее часто в литературе встречаются следующие способы такой формализации:

- Коэффициент Жаккара (Jaccard, 1901), равняющийся площади пересечения множеств Q и R , отнесенной к площади их объединения:

$$K_{\text{Jaccard}}(Q, R) \stackrel{\text{def}}{=} \frac{S(Q \cap R)}{S(Q \cup R)}. \quad (1)$$

Этот коэффициент является целевым, например, на конкурсе по анализу и распознаванию документов “ICDAR” (Zhukovsky et al., 2015).

- Метрика Хаусдорфа (Huttenlocher, 1993) и ее многочисленные модификации (Ефимов, Новиков, 2016; Dubuisson, Jain, 1994; Коваленко, Шемякина, 2018; Fresche, 1906). Она равняется наибольшему расстоянию от точек одного множества до соответствующих им ближайших точек второго множества:

$$d_{\text{Hausdorff}}(Q, R) \stackrel{\text{def}}{=} \max \left\{ \max_{q \in Q} \min_{r \in R} \|q - r\|_2, \max_{r \in R} \min_{q \in Q} \|r - q\|_2 \right\}. \quad (2)$$

Например, в работе (Dong-Gyu, 1999) метрика Хаусдорфа используется для проективного сопоставления объектов, в (Orrite, Herrero, 2004) — для проективного выравнивания частично заслоненных контуров, а в статье (Jesorsky, Kirchberg, 2001) — для робастной детекции лиц.

Рассмотрим две важные модификации метрики Хаусдорфа:

- Критерий максимальной невязки (Коваленко, Шемякина, 2018), который отличается от

метрики Хаусдорфа тем, что в нем вместо ближайшей точки для данной точки выбирается фактически соответствующая ей точка:

$$L_\infty(\hat{H}, H; R) \stackrel{\text{def}}{=} \max \left\{ \max_{q \in Q} \|q - \hat{H}\hat{H}^{-1}(q)\|_2, \max_{r \in R} \|r - \hat{H}\hat{H}^{-1}(r)\|_2 \right\}.$$

Заметим, что аргументы под первым максимумом равны, поэтому

$$L_\infty(\hat{H}, H; R) = \max_{r \in R} \|r - \hat{H}\hat{H}^{-1}(r)\|_2. \quad (3)$$

Этот критерий использовался в качестве точности нормализации фотографии документа перед распознаванием его текстового содержания (Коноваленко, Шемякина, 2018).

– Расстояние Фреше (Frechet, 1906) и его разновидности. Оно определяется следующим образом:

$$F(Q, R) = \inf_{\alpha, \beta} \max_{t \in [0,1]} \|\delta Q(\alpha(t)), \delta R(\beta(t))\|_2, \quad (4)$$

где δQ и δR – непрерывные биекции множества $[0,1]$ в границы множеств Q и R соответственно, а α, β – непрерывные монотонные сюръекции множества $[0, 1]$ в себя (репараметризация). Расстояние Фреше применялось, например, для измерения степени близости двух контуров после проективного выравнивания в задаче проективно-инвариантного распознавания плоских замкнутых контуров (Притула и др., 2014), для оценки качества сегментации изображений (Березский, Березская, 2015), для оценки качества сопоставления карт (Wei et al., 2013).

• Наиболее часто точность проективной нормализации рассчитывается как среднеквадратичная невязка координат:

$$L_2(\hat{H}, H; R) \stackrel{\text{def}}{=} \sqrt{\frac{1}{S(R)} \int_R \|r - \hat{H}\hat{H}^{-1}(r)\|_2^2 dr}. \quad (5)$$

Это можно видеть в работе (Hsu, Sawhney, 1998) в задаче автоматического исправления дисторсии, вызванной кривизной линзы и движением камеры, в статье (Chen, 2002), где решалась задача проективного выравнивания изображения, созданного при помощи матрицы проекторов, и во многих других работах, например (Baltzopoulos, 1995; Stein, 1997; Mallon, 2005).

Разделим эти критерии на две группы: $\{K_{\text{Jaccard}}, d_{\text{Hausdorff}}, F\}$ и $\{L_2, L_\infty\}$. Первая группа сравнивает множества Q и R , не учитывая попарные соответствия их элементов. Это приводит к тому, что существуют такие $\hat{H} \neq H$, для которых критерии возвращают значения, соответствующие максимальной точности: $d_{\text{Hausdorff}} = F = 0$, $K_{\text{Jaccard}} = 1$. Критерии второй группы учитывают попарные соответствия элементов Q и R и в случае $\hat{H} \neq H$ не

могут вернуть значения, соответствующие максимальной точности: $L_2 = L_\infty = 0$. В данной работе мы будем рассматривать только критерии второй группы.

КРИТЕРИИ ТОЧНОСТИ АФФИННОЙ АППРОКСИМАЦИИ ПРОЕКТИВНОГО ПРЕОБРАЗОВАНИЯ

Теперь будем считать, что преобразование H известно на практике, и вместо проективной оценки $\hat{H}$ проективного преобразования H будем рассматривать его аппроксимацию – аффинное преобразование A . Таким образом, мы адаптируем критерии точности проективного исправления для описания точности аффинной аппроксимации проективного преобразования:

$$L_2(A, H; R) = \sqrt{\frac{1}{S(R)} \int_R \|r - AH^{-1}(r)\|_2^2 dr}, \quad (6)$$

$$L_\infty(A, H; R) = \max_{r \in R} \|r - AH^{-1}(r)\|_2. \quad (7)$$

Зададим преобразование A матрицей $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ следующим образом:

$$A \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) \stackrel{\text{def}}{=} A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}. \quad (8)$$

Тогда:

$$L_2(A, H; R) = L_2(A, H; R) = \sqrt{\frac{1}{S(R)} \int_R \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2^2 dr}, \quad (9)$$

$$L_\infty(A, H; R) = L_\infty(A, H; R) = \max_{r \in R} \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2. \quad (10)$$

ОПТИМАЛЬНЫЕ АФФИННЫЕ АППРОКСИМАЦИИ ПРОЕКТИВНОГО ПРЕОБРАЗОВАНИЯ

Задача поиска оптимальной в смысле L_2 аффинной аппроксимации проективного преобразования H относительно области интереса R формулируется следующим образом:

$$A_{L_2}^* \stackrel{\text{def}}{=} \arg \min_A L_2(A, H; R) = \arg \min_A \sqrt{\frac{1}{S(R)} \int_R \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2^2 dr},$$

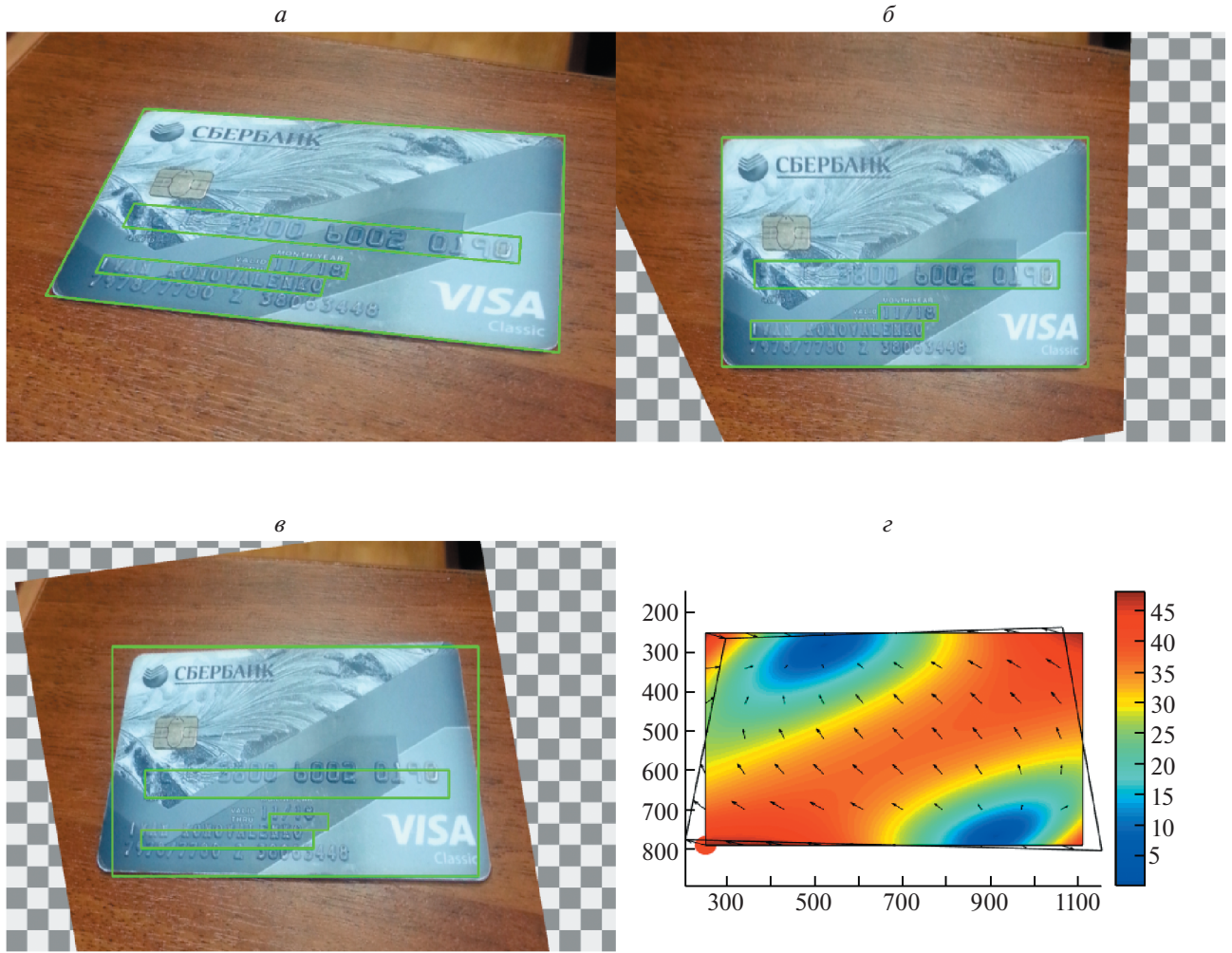


Рис. 1. *а* – исходное изображение; *б* – исходное изображение, исправленное проективным преобразованием H ; *в* – исходное изображение, исправленное оптимальной аффинной аппроксимацией $A_{L_\infty}^*$ преобразования H ; *г* – Область интереса R , векторное поле смещений $\left[A_{L_\infty}^* H^{-1}(r) - r \right]$, $r \in R$ и его ℓ_2 норма – невязка (показана цветом), положение максимальной невязки выделено красной точкой.

$$A_{L_2}^* = \arg \min_A \int_R \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2^2 dr, \quad (11)$$

а в смысле L_∞ – следующим:

$$A_{L_\infty}^* \stackrel{\text{def}}{=} \arg \min_A L_\infty(A, H; R),$$

$$A_{L_\infty}^* = \arg \min_A \max_{r \in R} \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2. \quad (12)$$

Докажем выпуклость задач (11) и (12).

Теорема 1. Задача (11) – выпуклая.

Доказательство. Задача называется выпуклой, если выпуклы допустимое множество и целевая

функция. Допустимое множество в нашем случае есть $\mathbb{R}^{2 \times 3}$, а значит, выпукло. Таким образом, осталось доказать выпуклость целевой функции по A . Выпишем ее

$$\int_R \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2^2 dr.$$

При каждом значении r подынтегральная функция $\left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2^2$ выпукла по A , а значит, выпуклый по A и интеграл (бесконечная сумма) таких функций, что и требовалось доказать.

Теорема 2. Задача (12) — выпуклая.

Доказательство. Отличие от предыдущего доказательства состоит в том, что тут требуется доказать выпуклость другой целевой функции:

$$\max_{r \in R} \left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2.$$

При каждом значении r функция $\left\| r - A \begin{bmatrix} H^{-1}(r) \\ 1 \end{bmatrix} \right\|_2$ выпукла по A , а значит, выпуклый по A и поточечный максимум таких функций, что и требовалось доказать.

Таким образом, обе задачи могут быть решены эффективно. Пример оптимальной аффинной аппроксимации проективного преобразования представлен на рис. 1.

ПРИМЕНЕНИЕ АФФИННОЙ АППРОКСИМАЦИИ ПРОЕКТИВНОГО ПРЕОБРАЗОВАНИЯ ДЛЯ УЛУЧШЕНИЯ БЫСТРОДЕЙСТВИЯ

С одной стороны, известно, что применение к растровому изображению аффинного преобразования занимает значительно меньше времени, чем применение проективного, и часто выполняется качественнее. С другой стороны, заменяя проективное преобразование на его быструю аффинную аппроксимацию, мы можем сильно потерять в геометрической точности. Поэтому мы предлагаем следующий подход, параметром которого является пороговая допустимая потеря точности T в смысле $L \in \{L_2, L_\infty\}$:

- По известным H и R вычисляем A_L^* .
- Если $L(A_L^*, H; R) \leq T$, то применяем к изображению преобразование A_L^* , иначе — H .

Таким образом, мы экономим вычислительные ресурсы только в тех случаях, когда проективное преобразование приближается аффинным преобразованием с высокой точностью.

ЗАКЛЮЧЕНИЕ

Показано, что для критериев точности, предложенных авторами в настоящей работе, задача оптимальной в смысле этих критериев аффинной аппроксимации является выпуклой, что делает возможным эффективный поиск глобального оптимума. Это позволяет исследовать точность возможной аппроксимации проективного преобразования аффинным и, если она достаточна, использовать вычислительно более эффективное аффинное преобразование изображения.

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 16-07-00616, № 17-29-03370 и № 17-29-03514.

СПИСОК ЛИТЕРАТУРЫ

- Балицкий А.М., Савчик А.В., Гафаров Р.Ф., Коноваленко И.А. О проективно инвариантных точках овала с выделенной внешней прямой. *Проблемы передачи информации*. 2017. Т. 53. № 3. С. 84–89.
- Березский О.Н., Березская К.М. Количественная оценка качества сегментации изображений на основе метрик. *Управляющие системы и машины*. 2015. № 6. С. 59–65.
- Болотова Ю.А., Спицын В.Г., Осина П.М. Обзор алгоритмов детектирования текстовых областей на изображениях и видеозаписях. *Компьютерная оптика*. 2017. Т. 41. № 3.
- Ефимов А.И., Новиков А.И. Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений. *Компьютерная оптика*. 2016. Т. 40. № 2.
- Коноваленко И.А., Шемякина Ю.А. Анализ величин ошибки при неточном проективном преобразовании четырехугольника. *Информационные технологии и нанотехнологии*. 2018. С. 1251–1260.
- Николаев П.П. Проективно инвариантное описание неплоских гладких фигур. 1. Предварительный анализ задачи. *Сенсорные системы*. 2016. Т. 30. № 4. С. 290–311.
- Питула Н.Е., Николаев П.П., Шешкус А.В. Сравнение двух алгоритмов проективно-инвариантного распознавания плоских замкнутых контуров с единственной вогнутостью. “Информационные технологии и системы”, Сборник трудов. 2014. С. 367–373.
- Савчик А.В., Николаев П.П. Теорема о пересечении T -и H -поляр. *Информационные процессы*. 2016. Т. 16. № 4. С. 430–443.
- Холопов И.С. Алгоритм коррекции проективных искажений при маловысотной съёмке. *Компьютерная оптика*. 2017. Т. 41. № 2.
- Чернов Т.С., Ильин Д.А., Безматерных П.В., Фараджев И.А., Карпенко С.М. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения. *Вестник Российского фонда фундаментальных исследований*. 2016. № 4. С. 55–71.
- Шемякина Ю.А. Использование точек и прямых для вычисления проективного преобразования по двум изображениям плоского объекта. *Информационные технологии и вычислительные системы*. 2017. № 3. С. 79–91.
- Шемякина Ю.А., Жуковский А.Е., Фараджев И.А. Исследование алгоритмов вычисления проективного преобразования в задаче наведения на планарный объект по особым точкам. *Искусственный интеллект и принятие решений*. 2017. № 1. С. 43–49.

- Alter T. 3D pose from 3 corresponding points under weak-perspective projection. *Technical report, Massachusetts inst. of tech. cambridge artificial intelligence lab*, 1992.
- Aradhye H., Myers G. *Method and apparatus for recognition of symbols in images of three-dimensional scenes*, US Patent 7,738,706. 2010.
- Baltzopoulos V. A videofluoroscopy method for optical distortion correction and measurement of knee-joint kinematics. *Clinical Biomechanics*. 1995. V. 10 (2). P. 85–92.
- Chen H., Sukthankar R., Wallace G., Li K. Scalable alignment of large-format multi-projector displays using camera homography trees. *Proceedings of the conference on Visualization'02*. 2002. P. 339–346.
- Darko P., Volker S., Leif K. Interactive image completion with perspective correction. *The Visual Computer*. 2006. V. 22 (9–11). P. 671–681.
- Dong-Gyu S., Oh-Kyu K., Rae-Hong P. Object matching algorithms using robust hausdorff distance measures. *IEEE Transactions on image processing*. 1999. V. 8 (3). P. 425–429.
- Dubuisson M., Jain A. A modified hausdorff distance for object matching. *Proceedings of the 12th IAPR International Conference*. 1994. V. 1. P. 566–568.
- Faugeras O. What can be seen in three dimensions with an uncalibrated stereo rig? *European conference on computer vision*. 1992. P. 563–578.
- Fréchet M. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo (1884–1940)*. 1906. V. 22 (1). P. 1–72.
- Gruen A. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*. 1985. V. 14 (3). P. 175–187.
- Heckbert P. Fundamentals of texture mapping and image warping. 1989. *UC Berkeley Master's thesis*. 86 p.
- Hsu S., Sawhney H. Influence of global constraints and lens distortion on pose and appearance recovery from a purely rotating camera. *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop*. 1998. P. 154–159.
- Huang J., Singh A., Ahuja N. Single image super-resolution from transformed self-exemplars. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015. P. 5197–5206.
- Huttenlocher D., Klanderman G., Rucklidge W. Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*. 1993. V. 15 (9). P. 850–863.
- Jaccard P. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bull Soc Vaudoise Sci Nat*. 1901. V. 37. P. 547–579.
- Jesorsky O., Kirchberg K.J., Frischholz R.W. Robust face detection using the hausdorff distance. *International Conference on Audio-and Video-Based Biometric Person Authentication*. 2001. P. 90–95.
- Kadir T., Zisserman A., Brady M. An affine invariant salient region detector. *European conference on computer vision*. 2004. P. 228–241.
- Karpenko S., Konovalenko I., Miller A., Miller B., Nikolaev D. UAV control on the basis of 3d landmark bearing-only observations. *Sensors*. 2015. V. 15 (12). P. 29802–29820.
- Kutulakos K.N., Vallino J. Affine object representations for calibration-free augmented reality. *Virtual Reality Annual International Symposium, 1996., Proceedings of the IEEE*. 1996. P. 25–36.
- Lorenz H., Döllner J. Real-time piecewise perspective projections. *GRAPP*. 2009. P. 147–155.
- Mallon J., Whelan P.F. Projective rectification from the fundamental matrix. *Image and Vision Computing*. 2005. V. 23 (7). P. 643–650.
- Mikolajczyk K., Schmid C. An affine invariant interest point detector. *European conference on computer vision*. 2002. P. 128–142.
- Mikolajczyk K., Schmid C. Scale & affine invariant interest point detectors. *International journal of computer vision*. 2004. V. 60 (1). P. 63–86.
- Morel J.-M., Yu G. Asift: A new framework for fully affine invariant image comparison. *SIAM journal on imaging sciences*. 2009. V. 2 (2). P. 438–469.
- Orrite C., Herrero J.E. Shape matching of partially occluded curves invariant under projective transformation. *Computer Vision and Image Understanding*. 2004. V. 93 (1). P. 34–64.
- Povolotskiy M.A., Kuznetsova E.G., Khanipov T.M. Russian license plate segmentation based on dynamic time warping. *31st European Conference on Modelling and Simulation (ECMS)*. 2017. P. 285–291.
- Skoryukina N., Chernov T., Bulatov K., Nikolaev D.P., Arlazarov V. Snapscreen: Tv-stream frame search with projectively distorted and noisy query. *Ninth International Conference on Machine Vision*. 2017. V. 10341. P. 103410Y.
- Skoryukina N., Shemiakina J., Arlazarov V., Faradjev I. Document localization algorithms based on feature points and straight lines. *Tenth International Conference on Machine Vision*. 2018. V. 10696. P. 106961H.
- Stein G.P. Lens distortion calibration using point correspondences. *Computer Vision and Pattern Recognition. Proceedings. IEEE Computer Society Conference*. 1997. P. 602–608.
- Tu-ichi O., Kiyoshi M., Toshiyuki S. Obtaining surface orientation from texels under perspective projection. *IJCAI*. 1981. P. 746–751.
- Wei H., Wang Y., Forman G., Zhu Y. Map matching by fréchet distance and global weight optimization. *Technical Paper, Departement of Computer Science and Engineering*. 2013. P. 19.
- Zhukovsky A.E., Arlazarov V.V., Postnikov V.V., Krivtsov V.E. Segments graph-based approach for smartphone document capture. *Eighth International Conference on Machine Vision*. 2015. V. 9875. P. 98750P.
- Zwicker M., Räsänen J., Botsch M., Dachsbacher C., Pauly M. Perspective accurate splatting. *Proceedings of Graphics interface 2004*. 2004. P. 247–254.

Optimal affine approximation of image projective transformation

I. A. Konovalenko^{a,b,#}, V. V. Kokhan^a, and D. P. Nikolaev^{a,b}

^a Institute for Information Transmission Problems of Russian Academy of Sciences (Kharkevich Institute) – IITP RAS
127051 Moscow, Bolshoy Karetny lane 19, build.1, Russia

^b Smart Engines Ltd., 117312 Moscow, 60-letiya Oktyabrya avenue, 9, Russia

[#] E-mail: konovalenko@iitp.ru

Replacing the projective transformation with a substantially simpler affine transformation occurs in many areas of technical vision. At the same time, the concept of accuracy of an affine approximation of a projective transformation is not formalized in the literature. This, in turn, leads to the absence of problem statements and theoretically sound methods for affine approximation of the projective transformation. This work aims to eliminate this gap. The authors proposed to use the root mean square (RMS) and the maximum pointwise discrepancy in the transformed image coordinates system as criteria for the accuracy of the affine approximation of projective transformation. Based on these criteria, the problem of finding the optimal affine approximations is formulated. The convexity of the obtained optimization problems is proved. A method for image transformation, using optimal affine approximations to save computational resources, is proposed.

Key words: homography, homography estimation accuracy, transformation affine approximation, linearization, convex analysis

REFERENCES

- Alter T. 3D pose from 3 corresponding points under weak-perspective projection. *Technical report, Massachusetts inst of tech cambridge artificial intelligence lab*, 1992.
- Aradhye H., Myers G. *Method and apparatus for recognition of symbols in images of three-dimensional scenes*, US Patent 7,738,706. 2010.
- Balickij A.M., Savchik A.V., Gafarov R.F., Konovalenko I.A. O proektivno invariantnyh tochkah ovala s vydelennoj vneshnej prjamoj [On projectively invariant points of an oval with a distinguished outer straight line] *Problemy peredachi informacii* [Problems of Information Transmission]. 2017. V. 53 (3). P. 84–89 (in Russian).
- Berezskij O.N., Berezskaja K.M. Kolichestvennaja ocenka kachestva segmentacii izobrazhenij na osnove metric [Quantitative assessment of the quality of image segmentation based on metrics] *Upravljajushhie sistemy i mashiny* [Control systems and machines]. 2015. № 6. P. 59–65 (in Russian).
- Bolotova Ju.A., Spicyn V.G., Osina P.M. Obzor algoritmov detektirovaniya tekstovyh oblastej na izobrazhenijah i videozapisjah [An overview of the algorithms for detecting text areas in images and video recordings] *Komp'yuternaja optika* [Computer optics]. 2017. V. 41 (3) (in Russian).
- Chen H., Sukthankar R., Wallace G., Li K. Scalable alignment of large-format multi-projector displays using camera homography trees. *Proceedings of the conference on Visualization'02*. 2002. P. 339–346.
- Chernov T.S., Il'in D.A., Bezmaternyh P.V., Faradzhev I.A., Karpenko S.M. Issledovanie metodov segmentacii izobrazhenij tekstovyh blokov dokumentov s pomoshh'ju algoritmov strukturnogo analiza i mashinnogo obucheniya [Study of methods for segmentation of images of text blocks of documents using algorithms for structural analysis and machine learning] *Vestnik Rossijskogo fonda fundamental'nyh issledovanij* [Vestnik RFFI]. 2016. № 4. P. 55–71 (in Russian).
- Darko P., Volker S., Leif K. Interactive image completion with perspective correction. *The Visual Computer*. 2006. V. 22 (9–11). P. 671–681.
- Dong-Gyu S., Oh-Kyu K., Rae-Hong P. Object matching algorithms using robust hausdorff distance measures. *IEEE Transactions on image processing*. 1999. V. 8 (3). P. 425–429.
- Dubuisson M., Jain A. A modified hausdorff distance for object matching. *Proceedings of the 12th IAPR International Conference*. 1994. V. 1. P. 566–568.
- Efimov A.I., Novikov A.I. Algoritm pojetapnogo utochenija proektivnogo preobrazovanija dlja sovmeshhenija izobrazhenij [Algorithm of phased refinement of projective transformation for image combining] *Komp'yuternaja optika* [Computer optics]. 2016. V. 40 (2) (in Russian).
- Faugeras O. What can be seen in three dimensions with an uncalibrated stereo rig? *European conference on computer vision*. 1992. P. 563–578.
- Fréchet M. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo (1884–1940)*. 1906. V. 22 (1). P. 1–72.
- Gruen A. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*. 1985. V. 14 (3). P. 175–187.
- Heckbert P. Fundamentals of texture mapping and image warping. *UC Berkeley Master's thesis*. 1989. 86 p.
- Holopov I. S. Algoritm korrekcii proektivnyh iskazhenij pri malovysothnoj s#jomke [Algorithm for the correction of projective distortion in low-altitude shooting] *Komp'yuternaja optika* [Computer optics]. 2017. V. 41 (2) (in Russian).
- Hsu S., Sawhney H. Influence of global constraints and lens distortion on pose and appearance recovery from a purely rotating camera. *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop*. 1998. P. 154–159.

- Huang J., Singh A., Ahuja N. Single image super-resolution from transformed self-exemplars. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015. P. 5197–5206.
- Huttenlocher D., Klanderman G., Rucklidge W. Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*. 1993. V. 15 (9). P. 850–863.
- Jaccard P. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bull Soc Vaudoise Sci Nat*. 1901. V. 37. P. 547–579.
- Jesorsky O., Kirchberg K.J., Frischholz R.W. Robust face detection using the hausdorff distance. *International Conference on Audio-and Video-Based Biometric Person Authentication*. 2001. P. 90–95.
- Kadir T., Zisserman A., Brady M. An affine invariant salient region detector. *European conference on computer vision*. 2004. P. 228–241.
- Karpenko S., Konovalenko I., Miller A., Miller B., Nikolaev D. UAV control on the basis of 3d landmark bearing-only observations. *Sensors*. 2015. V. 15 (12). P. 29802–29820.
- Konovalenko I.A., Shemjakina Ju.A. Analiz velichin oshibki pri netochnom proektivnom preobrazovanii chetyrehugol'nika [Error values analysis for inaccurate projective transformation of a quadrangle]. *Informacionnye tehnologii i nanotekhnologii* [information technology and nanotechnology]. 2018. P. 1251–1260 (in Russian).
- Kutulakos K.N., Vallino J. Affine object representations for calibration-free augmented reality. *Virtual Reality Annual International Symposium. Proceedings of the IEEE*. 1996. P. 25–36.
- Lorenz H., Döllner J. Real-time piecewise perspective projections. *GRAPP*. 2009. P. 147–155.
- Mallon J., Whelan P.F. Projective rectification from the fundamental matrix. *Image and Vision Computing*. 2005. V. 23 (7). P. 643–650.
- Mikolajczyk K., Schmid C. An affine invariant interest point detector. *European conference on computer vision*. 2002. P. 128–142.
- Mikolajczyk K., Schmid C. Scale & affine invariant interest point detectors. *International journal of computer vision*. 2004. V. 60 (1). P. 63–86.
- Morel J-M., Yu G. Asift: A new framework for fully affine invariant image comparison. *SIAM journal on imaging sciences*. 2009. V. 2 (2). P. 438–469, 2009.
- Nikolaev P.P. Proektivno invariantnoe opisanie neploskih gladkih figur. 1. Predvaritel'nyj analiz zadachi [Projectively invariant description of non-planar smooth figures. 1. Preliminary analysis of the problem] *Sensornye sistemy* [Sensory systems]. 2016. V. 30 (4). P. 290–311 (in Russian).
- Orrite C., Herrero J.E. Shape matching of partially occluded curves invariant under projective transformation. *Computer Vision and Image Understanding*. 2004. V. 93 (1). P. 34–64.
- Povolotskiy M.A., Kuznetsova E.G., Khanipov T.M. Russian license plate segmentation based on dynamic time warping. *31st European Conference on Modelling and Simulation (ECMS)*. 2017. P. 285–291.
- Pritula N.E., Nikolaev P.P., Sheshkus A.V. Sravnenie dvuh algoritmov proektivno-invariantnogo raspoznavaniya ploskih zamknutyh konturov s edinstvennoj vognutost'ju [Comparison of two algorithms of projectively invariant recognition of flat closed contours with a single concavity] *Informacionnye tehnologii i sistemy-2014* [Information technologies and systems-2014], Almannac. ISBN 978-5-901158-25-8 C. 367–373 P. 367–373 (in Russian).
- Savchik A.V., Nikolaev P.P. Teorema o peresechenii T-i H-poljar [The Theorem of T- and H- Polars Intersections Count] *Informacionnye process* [Information processes]. 2016. V. 16(4). P. 430–443 (in Russian).
- Shemjakina Ju. A. Ispol'zovanie toчек i prjamyh dlja vychislenija proektivnogo preobrazovanija po dvum izobrazhenijam ploskogo ob'ekta [Using points and lines to calculate the projective transformation from two images of a flat object] *Informacionnye tehnologii i vychislitel'nye sistemy* [information technology and computing systems]. 2017. No. 3. P. 79–91. (in Russian).
- Shemjakina Ju. A., Zhukovskij A. E., Faradzhev I. A. Issledovanie algoritmov vychislenija proektivnogo preobrazovanija v zadache navedenija na planarnyj ob'ekt po osobym tochkam [The Research of the Algorithms of a Projective Transformation Calculation in the Problem of Planar Object Targeting by Feature Points] *Iskusstvennyj intellekt i prinjatje reshenij* [Artificial intelligence and decision making]. 2017. No. 1. P. 43–49 (in Russian).
- Skoryukina N., Chernov T., Bulatov K., Nikolaev D.P., Arlazarov V. Screenshot: Tv-stream frame search with projectively distorted and noisy query. *Ninth International Conference on Machine Vision*. 2017. V. 10341. P. 103410Y.
- Skoryukina N., Shemiakina J., Arlazarov V., Faradjev I. Document localization algorithms based on feature points and straight lines. *Tenth International Conference on Machine Vision*. 2018. V. 10696. P. 106961H.
- Stein G.P. Lens distortion calibration using point correspondences. *Computer Vision and Pattern Recognition. Proceedings. IEEE Computer Society Conference*. 1997. P. 602–608.
- Tu-ichi O., Kiyoshi M., Toshiyuki S. Obtaining surface orientation from texels under perspective projection. *IJCAI*. 1981. P. 746–751.
- V Baltzopoulos. A videofluoroscopy method for optical distortion correction and measurement of knee-joint kinematics. *Clinical Biomechanics*. 1995. V. 10 (2). P. 85–92.
- Wei H., Wang Y., Forman G., Zhu Y. Map matching by fréchet distance and global weight optimization. *Technical Paper, Departement of Computer Science and Engineering*. 2013. P. 19.
- Zhukovsky A.E., Arlazarov V.V., Postnikov V.V., Krivtsov V.E. Segments graph-based approach for smartphone document capture. *Eighth International Conference on Machine Vision*. 2015. V. 9875. P. 98750P.
- Zwicker M., Räsänen J., Botsch M., Dachsbacher C., Pauly M. Perspective accurate splatting. *Proceedings of Graphics interface* 2004. 2004. P. 247–254.