



УДК 004.932.2

© 2022 г. **А.В. Зацерковный**

(Тихоокеанский океанологический институт им. В.И. Ильичева,
Владивосток),

Е.А. Нурминский, д-р физ.-мат. наук

(Дальневосточный федеральный университет, Владивосток)

ОЦЕНКА И ИСПОЛЬЗОВАНИЕ ПАРАМЕТРОВ ПУБЛИЧНЫХ ВИДЕОКАМЕР ДЛЯ НАБЛЮДЕНИЯ ЗА АВТОМОБИЛЬНЫМ ДВИЖЕНИЕМ*

При использовании общедоступных видеокамер наружного наблюдения для сбора адекватной информации о параметрах дорожного движения необходимо учитывать характеристики камер, их размещение и другую информацию, которая непосредственно недоступна. Эта работа описывает разработанную авторами технологию обработки видеопотока изображений, позволяющую в значительной степени избавиться от перспективных и радиальных искажений, что даст возможность повысить точность распознавания типов транспортных средств, определения плотности автомобильного потока и его скоростных характеристик.

Ключевые слова: машинное зрение, калибровка камер, компенсация искажений.

DOI: 10.22250/18142400_2022_73_3_93

Введение

Съем достоверных и достаточно точных данных о количественных характеристиках транспортных потоков критически важен. Эти данные должны покрывать весь временной интервал существенного трафика (в современных условиях значительно больше 12 часов в день), для учета сезонной и погодной изменчивости, измерения должны производиться на постоянной основе и в как можно большем количестве мест. Публично доступные каме-

* Исследование выполнено при финансовой поддержке Министерства науки и высшего образования РФ (проект № 075-02-2022-880).

ры видеонаблюдения могут быть ценным источником таких данных для исследователей. Пример анализа параметров транспортных потоков по данным веб-камеры в части сцены, практически не подверженной перспективному и радиальному искажениям, предложен в [1], где описан подход, позволяющий обобщить такой анализ на сцены с существенными перспективным и радиальным искажениями. В работе по данным, которые несложно собрать для публичных камер стороннего оператора, оцениваются параметры камеры. Параметры камеры используются для получения отображения, позволяющего оценивать метрические расстояния в пространстве, соответствующие расстояниям между пикселями на изображении. Этой функции достаточно, чтобы применять к видеосцене анализ, описанный в [1].

Системы координат, модели

Для локализации в пространстве выберем ENU систему координат F_{enu} (East, Nord, Up), координаты точки P в которой будем обозначать: $P = (x_e, y_e, z_e)$.

Камера строит изображение на плоскости фотосенсора (рис. 1, 2). Изображение представляет собой матрицу $N \times M$ из пикселей (N столбцов, M строк). Положение пикселя на матрице описывается в системе координат изображения F_{im} . Ось U ориентирована вдоль строк матрицы (нумерация слева направо), ось V направлена вдоль столбцов (нумерация сверху вниз). Плоскость сенсора ориентирована по камере. Вещественные координаты точки P_{im} в F_{im} обозначим: $P_{im} = (u, v)$. Для локализации пикселя нужны целочисленные индексы, которые можно получать из координат (u, v) округлением до ближайшего целого. Пусть w, h – ширина и высота пикселя соответственно. Для полной ширины и высоты изображения справедливы равенства: $W = Nw, H = Mh$.

К камере привязана система координат камеры F_c (рис. 1). Начало расположено в точке O – центре перспективной проекции. Координаты точки O в F_{enu} обозначим (O_{xe}, O_{ye}, O_{ze}) . Ось Z_c ориентирована вдоль оптической оси объектива, X_c направлена параллельно оси U , Y_c параллельна оси V . Ось X_c называют горизонтом камеры, а ее отклонение от горизонтальной плоскости – завалом горизонта. F_{enu} и F_c – метрические системы координат. В F_{im} вещественные координаты определяют индексы пикселей в матрице. Чтобы описать координатами u, v , равными в метрах расстояниям \tilde{u}, \tilde{v} , требуется соблюдение равенства $\tilde{u} = wu = hv \tilde{v}$.

Камера формирует перспективную проекцию пространственной сцены в плоскости изображения, используя преобразование, которое описывают моделью камеры-обскуры (pinhole camera, рис. 1, 2 и [2]). Изображение точки $P_{im} = (u, v)$ в модели является результатом преобразования координат ориги-

нала P , заданных в F_{enu} . Преобразование описывается матрицами \mathbf{A} , $[\mathbf{R}|t]$:

$$\mathbf{A} = \begin{bmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

$$[\mathbf{R}|t] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}. \quad (2)$$

\mathbf{A} называют матрицей внутренних (intrinsic) параметров камеры. Матрица $[\mathbf{R}|t]$ определяет переход от ENU координат к координатам камеры (ее содержимое определяется местом установки и ориентацией камеры):

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = [\mathbf{R}|t] \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}, \quad \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = -\mathbf{R} \begin{bmatrix} O_{xe} \\ O_{ye} \\ O_{ze} \end{bmatrix}. \quad (3)$$

\mathbf{R} является матрицей вращения, t описывает сдвиг в пространстве из начала координат F_{enu} в O .

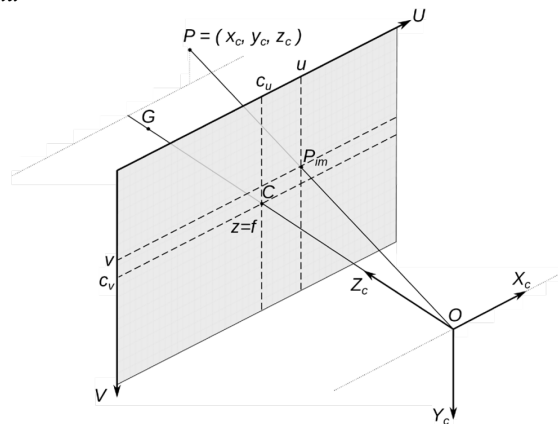


Рис. 1. Системы координат F_{im} , F_c , изображение P_{im} трехмерной точки P . Ось Z_c направлена вдоль оптической оси камеры. Изображение формируется на сенсоре за отверстием O в плоскости $z = -f$, в иллюстрациях удобнее эквивалентное "виртуальное" изображение в плоскости $z = f$.

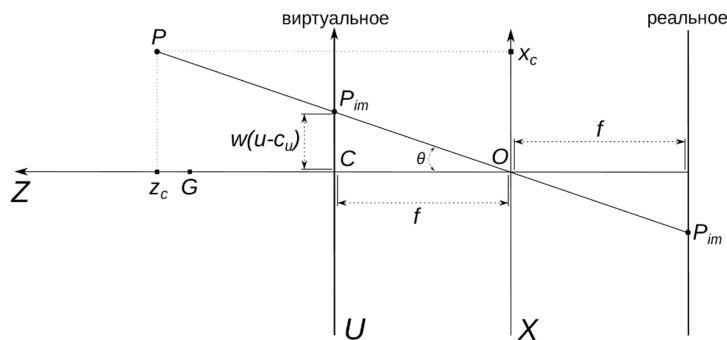


Рис. 2. Проекция сцены на плоскость XZ_c , показывающая положение виртуального и реального изображений относительно апертуры O . Здесь x_c, z_c – координаты точки P в F_c ; u, c_u – координаты изображения точек P и C в F_{im} .

Матрица \mathbf{A} содержит координаты точки $C = (c_u, c_v)$ в F_{im} и отношения фокальной длины камеры к ширине и высоте пикселя

$$f_u = f / \omega, \quad f_v = f / h, \quad (4)$$

Точка C (оптический центр изображения, principal point) расположена на пересечении оптической оси камеры с изображением и является изображением пространственной точки G . Координаты (c_u, c_v) обычно указывают на центр изображения (т.е. для Full HD камеры 1920x1080 $c_u = 960, c_v = 540$).

Учитывая подобие проекций треугольников (O, P_{im}, C) , $(O, P, (z_c, 0, 0))$ на плоскости ZX_c и ZY_c (рис. 2), справедливы следующие равенства:

$$\begin{aligned} x_c / z_c &= \omega(u - c_u) / f = \tan \theta_x, \\ y_c / z_c &= h(v - c_v) / f = \tan \theta_y. \end{aligned} \quad (5)$$

Из (5) следует, что координаты пикселя в F_{im} можно выразить через координаты его прообраза в F_c (рис. 2):

$$\mathbf{P}_{im}^h = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_c / z_c \\ y_c / z_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_u x_c / z_c + c_u \\ f_v y_c / z_c + c_v \\ 1 \end{bmatrix}, \quad (6)$$

где \mathbf{P}_{im}^h – расширение P_{im} в гомогенных координатах. Из (4), (5) видно, что при $f_u \neq 0, f_v \neq 0$ для заданного пикселя (u, v) можно определить значения $x_c / z_c, y_c / z_c$. Однако для определения пространственных координат (x_c, y_c, z_c) прообраза пикселя (u, v) нужно знать $z_c(u, v)$. Множество значений $z_c(u, v)$ называют картой глубины изображения, для ее составления нужна дополнительная информация о наблюдаемой сцене.

Для повышения точности модель камеры обскура дополняют моделью радиального искажения, вносимого линзами объектива. Влияние радиального искажения на широкоугольных камерах хорошо заметно, особенно на краях изображений. Популярная квадратичная модель радиального искажения [2, 7] предлагает следующую модификацию к (6):

$$r^2 = (x_c^2 + y_c^2) / z_c^2, \quad \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} x_c / z_c \\ y_c / z_c \end{bmatrix}, \quad (7)$$

$$\mathbf{P}_{im}^h = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \hat{x} \\ \hat{y} \\ 1 \end{bmatrix} = \begin{bmatrix} f_u \hat{x} + c_u \\ f_v \hat{y} + c_v \\ 1 \end{bmatrix}, \quad (8)$$

где k_1, k_2 – коэффициенты радиального искажения.

Оценка параметров публичной камеры

Теперь можно перечислить параметры, определяющие преобразование пространственных точек в поле зрения камеры в точки на изображении: мат-

рица вращения R (2), (3), координаты точки O в F_{enu} (3), внутренние параметры камеры f_u, f_v, c_u, c_v (1), (4), коэффициенты радиального искажения k_1, k_2 (7), (8). Процесс определения значений параметров преобразования по заданной информации о положении точек $\{P_{enu}\}$ в пространстве и положении изображений этих точек $\{P_{im}\}$ на множестве снимков называют калибровкой камеры.

Пространственные ENU координаты множества точек можно довольно точно оценить с помощью трехмерных лидаров либо геодезических инструментов. Когда точные инструменты недоступны, можно использовать оценку глубины по стерео/мультиснимкам либо глобальные координаты точек, полученные, например, с карт и от GNSS сенсоров. Последние варианты широко доступны, но не обеспечивают точность на отдельных измерениях. Глобальные координаты можно преобразовать в ENU координаты [6]. Множество $\{P_{enu}\}$ называют облаком точек (point cloud). Существует арсенал инструментов для работы с облаком точек и для локализации изображений пространственных точек.

Процедура калибровки исследуется давно. Разработаны ее успешные реализации по набору подходящих изображений с камеры [2, 3, 7, 8]. Готовые процедуры возвращают оценку внутренних параметров камеры, оценки положения камеры относительно специального образца. После установки камеры для реальных наблюдений данные о положении камеры по отношению к калибровочному образцу обычно бесполезны. Если на камере применяется объектив с оптическим зумом и на месте потребовалось поменять его настройки, фокусное расстояние объектива меняется, калибровку приходится делать заново. Когда используется публичная камера, установленная оператором на здании или мачте, готовую функцию вроде *calibrateCamera* библиотеки OpenCV [2] применить сложно. Обычно оператор не выкладывает внутренние параметры и коэффициенты искажения камеры, но эта информация необходима для применения многих алгоритмов машинного зрения. Приходится искать другие пути калибровки.

Иногда удобнее отделить оценку расположения камеры (параметры R, O) от вычисления остальных параметров. Информация о разрешении $N \times M$ камеры доступна с данными. Как уже отмечалось, обычно

$$c_u = N / 2, c_v = M / 2. \quad (9)$$

Многие фотокамеры (в том числе мобильные) включают в файл изображения EXIF метаданные об условиях съемки. Среди EXIF данных нередко есть фокальная длина f . Если удастся получить внутренние параметры камеры и коэффициенты искажения, то R, O можно вычислить по множествам

$\{P_{enu}\}, \{P_{im}\}$ с помощью готовых реализаций алгоритма pose computation (PnP, [2, 9]). Когда EXIF данные камеры не помогают, значения f_u, f_v можно оценить по наборам изображений точек и координат оригиналов, используя равенства (5).

Нередко место установки публичной камеры определяется визуально. В таком случае ее ENU координаты (координаты точки O) можно оценить с помощью одного из указанных ранее подходов. Ориентацию камеры можно установить, используя координаты прообраза оптического центра C (точки G , рис. 1) и наблюдаемый на изображении наклон прямых, которые должны быть вертикальными или горизонтальными в пространстве. Такие прямые можно строить на изображении искусственно.

Рассмотрим пример оценки параметров конкретной публичной камеры в варианте, когда доступны только изображения с камеры и небольшой набор пространственных измерений ограниченной точности (рис. 3). Данная камера является хорошим примером, так как ее видео содержит видимые перспективное и радиальное искажения. На изображениях с камеры есть линия, по которой заметен завал горизонта (стена, в основании которой находится точка p_2). Камера отдает видео в формате FullHD ($M = 1080, N = 1920$).



Рис.3. Поле зрения одной из публичных дорожных камер Владивостока.

Фиксируем начало координат F_{enu} в точке p_0 (рис. 3), оценим ENU координаты камеры O , оригинала оптического центра G (рис. 1, 3) и точек p_1, p_2 , лежащих, соответственно, в плоскостях $y_c = 0$ и $x_c = 0$ системы координат камеры F_c . Начнем с оценки глобальных координат этих точек (вынужденная мера в отсутствии геодезического оборудования).

Используя карты и GNSS сенсоры, мы получили следующие значения:

Имя	Широта	Долгота	Высота	u	v
p_0	43.175553°	131.917725°	56		
G	43.176295°	131.918380°	57	960	540
O	43.176934°	131.917912°	98		
p_1	43.176033°	131.917937°	57	1534	540
p_2	43.175828°	131.918728°	52	960	353

Широта и долгота записаны в градусах, высота над уровнем моря в метрах, в последних столбцах записаны координаты изображений точек (пикселей) в F_{im} (**u** – столбец, **v** – строка). Точки можно найти на спутниковом слое онлайн карт [5], используя широту и долготу в запросе. Преобразуем глобальные координаты в координаты F_{enu} [6]. Теперь можно оценить углы между оптической осью OG и векторами Op_1 и Op_2 (рис. 1, 2):

$$\begin{aligned}\theta_x &= \arccos\left(\overline{OG} \cdot \overline{Op_1} / (|\overline{OG}| |\overline{Op_1}|)\right), \\ \theta_y &= \arccos\left(\overline{OG} \cdot \overline{Op_2} / (|\overline{OG}| |\overline{Op_2}|)\right).\end{aligned}\quad (10)$$

Обозначим координаты изображения точки p_1 в F_{im} (u_1, v_1) и p_2 соответственно (u_2, v_2). Из (5) следует, что если влиянием радиального искажения на положение изображений p_1, p_2 можно пренебречь (т.е. радиальное искажение умеренное и изображения точек не слишком удалены от C), то

$$f_u = (u_1 - c_u) / \tan \theta_x, \quad f_v = (u_2 - c_v) / \tan \theta_y. \quad (11)$$

Это выражение может дать качественное начальное приближение для коэффициентов f_u, f_v , обеспечив оценку матрицы внутренних параметров камеры A . Используя (10), (11) для точек p_1, p_2 примера, получим $f_u = 1166.2$, $f_v = 1241.55$. Размеры пикселей камер видеонаблюдения обычно близки к квадратным. При разнице в единицах процентов значения f_u, f_v можно положить равными (используем среднее), получим $f_u = f_v = 1204.4$.

Наличие матрицы A и набора $\{P_{enu}\}, \{P_{im}\}$ превращают (8) в набор линейных уравнений для коэффициентов k_1, k_2 . При этом можно оценить k_1, k_2 и компенсировать радиальное искажение основной части изображения [7]. Положим $k_1 = -0.24$, $k_2 = 0$ и применим функцию $undistort()$ библиотеки OpenCV [3] для устранения радиального искажения k_1, k_2 камеры, характеризуемой матрицей A .

В результате (рис. 4) радиальное искажение линий, которые должны быть прямыми, практически исчезло. Поле зрения камеры уменьшилось, точка C осталась на месте, точки p_1, p_2 сдвинулись далее вдоль прямых $v = c_v$, $u = c_u$. Значения f_u, f_v можно пересчитать, но радиальное искажение – лишь одна из причин ошибок, есть еще погрешности в оценке пространственных координат. Поэтому проведем дополнительные перекрестные проверки и компенсируем значения f_u, f_v по их результатам позднее, если потребуется.



Рис. 4. Изображение с камеры после компенсации радиального искажения.

Для определения ориентации камеры используем точку установки камеры O и точку G на оптической оси камеры (прообраз C), заданные в координатах F_{enu} (рис. 1). Орт

$$\mathbf{e}_{zc} = \overline{OG} / |\overline{OG}| = (e_x^{zc}, e_y^{zc}, e_z^{zc})_{enu} \quad (12)$$

дает направление оси Z_c координат F_c . Вектор \mathbf{e}_{zc} и центр координат камеры O определяют плоскость точек x , для которых вектор Ox перпендикулярен \mathbf{e}_{zc} . В этой плоскости должны лежать оси X_c, Y_c системы координат камеры.

Единичный вектор $\mathbf{e}_d = (0,0,1)_{enu}$ смотрит вниз. Если горизонт камеры не завален, ось X_c системы координат F_v такой камеры будет перпендикулярна вектору \mathbf{e}_d , так как ось должна быть параллельна горизонтальной плоскости. Мы можем найти направления осей Y_c, Z_c системы F_v :

$$\begin{aligned} \mathbf{n}_{yc} &= \mathbf{e}_d - (\mathbf{e}_d \cdot \mathbf{e}_{zc})\mathbf{e}_{zc} = \mathbf{e}_d + e_z^{zc}\mathbf{e}_{zc}, \\ \mathbf{e}_{yc} &= \mathbf{n}_{yc} / |\mathbf{n}_{yc}|, \quad \mathbf{e}_{xc} = \mathbf{e}_{yc} \times \mathbf{e}_{zc}. \end{aligned} \quad (13)$$

Векторы $(\mathbf{e}_{xc}, \mathbf{e}_{yc}, \mathbf{e}_{zc})$ образуют ортонормированный базис, который позволяет построить матрицу вращения R_{v2e} для перехода от F_v к F_{enu} :

$$\mathbf{R}_{v2e} = \begin{bmatrix} e_x^{xc} & e_x^{yc} & e_x^{zc} \\ e_y^{xc} & e_y^{yc} & e_y^{zc} \\ e_z^{xc} & e_z^{yc} & e_z^{zc} \end{bmatrix} = [\mathbf{e}_{xc} \quad \mathbf{e}_{yc} \quad \mathbf{e}_{zc}]. \quad (14)$$

По изображению можно оценить угол завала и повернуть плоскость с осями X_c, Y_c системы F_v вокруг оптической оси на этот угол, чтобы результирующая система координат F_c соответствовала реальной ориентации камеры.

В результате поворота изображения вокруг оптической оси камеры на

4.1° удалось добиться вертикальности нужных линий. Для перехода из координат F_c камеры в F_{enu} можно сначала перейти из F_c в F_v поворотом на угол $\gamma = 4.1^\circ$ вокруг оптической оси камеры OZ с помощью матрицы вращения:

$$\Gamma = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (15)$$

Переход из F_c в F_{enu} (без учета смещения начала координат) можно описать как комбинацию вращений матрицей (которая также является матрицей вращения)

$$\mathbf{R}_{c2e} = \mathbf{R}_{v2e} \Gamma. \quad (16)$$

Обратный переход от F_{enu} в координаты камеры F_c задается матрицей, которую мы уже обозначили \mathbf{R} (2):

$$\mathbf{R} = \mathbf{R}_{c2e}^{-1} = \mathbf{R}_{c2e}^T \quad (17)$$

и смещением t , которое в координатах камеры характеризует переход от начала координат F_{enu} в точку установки камеры. Смещение t часто удобнее выразить через координаты точки O , заданные в F_{enu} (3) и значения координат O_{enu} . Чтобы преобразовать координаты точки P_{enu} , заданной в F_{enu} в координаты P_c , для системы камеры F_c используем следующие выражения:

$$P_c = R(P_{enu} - O_{enu}) = RP_{enu} + t. \quad (18)$$

Использование параметров камеры

Далее рассматриваем изображение после компенсации радиального искажения, полагая применимость уравнений (6) и заданные f_u, f_v, c_u, c_v . Используем пространственные координаты точек, заданные в F_c (полученные после преобразования из F_{enu} , для чего необходимы R, O_{enu}).

Применим параметры камеры для получения карты глубины $z_c(u, v)$ пикселей, изображающих поверхность проезжей части дороги (u, v – координаты пикселя в F_{im}). Рассмотрим полосы движения в направлении, ближайшем к камере. Выберем несколько точек в этой области для аппроксимации поверхности дороги плоскостью, оценим их глобальные координаты:

№ точки	Широта	Долгота	Высота
1	43.176500°	131.918103°	59.36
2	43.176442°	131.918310°	59
3	43.176532°	131.918362°	59.3
4	43.176329°	131.918286°	58.6
5	43.176139°	131.917970°	57.
6	43.176093°	131.918187°	56.9
7	43.175553°	131.917725°	56
8	43.175495°	131.917991°	55.9
9	43.175794°	131.917845°	56

Точки расположены по краю проезжей части, одна из этих точек – p_0 – начало координат F_{enu} . Преобразуем глобальные координаты в F_{enu} , следуя [6], затем в F_c , следуя (18). Рассматриваемый участок дороги является практически плоским, с уклоном. Аппроксимируем его с помощью метода наименьших квадратов. Плоскость определяется уравнением

$$p_x x_c + p_y y_c + z_c = p_z, \mathbf{p} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}. \quad (19)$$

Представим данные о точках на дороге, полученные из последней таблицы, в виде матрицы \mathbf{D} и вектора \mathbf{E} :

$$\mathbf{D} = \begin{bmatrix} x_c^1 & y_c^1 & -1 \\ x_c^2 & y_c^2 & -1 \\ \dots & \dots & \dots \\ x_c^9 & y_c^9 & -1 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} -z_c^1 \\ -z_c^2 \\ \dots \\ -z_c^9 \end{bmatrix}. \quad (20)$$

Требуется найти $\text{argmin}_p(\mathbf{Dp} - \mathbf{E})$, явное решение задачи известно:

$$\mathbf{p} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{E} = (-0.20316, 2.04433, 86.99813). \quad (21)$$

Обозначим Π плоскость, определяемую (19), (21). Для точки $(x_c, y_c, z_c) \in \Pi$ изображенной в пикселе (u, v) , с учетом (6), имеем:

$$z_c = p_z - p_x x_c - p_y y_c, \quad x_c / z_c = (u - c_u) / f_u, \quad y_c / z_c = (v - c_v) / f_v. \quad (22)$$

Если $u \neq c_u$ и $v \neq c_v$ (из (22) следует, что $x_c \neq 0$ и $y_c \neq 0$), то

$$p_z - p_x x_c - p_y y_c = x_c f_u / (u - c_u), \quad p_z - p_x x_c - p_y y_c = y_c f_v / (v - c_v) \quad (23)$$

и значит есть линейные уравнения, позволяющие выразить x_c, y_c через u, v . Положим

$$a(u) = f_u / (u - c_u) + p_x, \quad b(v) = f_v / (v - c_v) + p_y. \quad (24)$$

Из (23) получим решения:

$$z_c(u, v) = p_z - p_x x_c(u, v) - p_y y_c(u, v), \quad (25)$$

$$\begin{bmatrix} x_c(u, v) \\ y_c(u, v) \end{bmatrix} = \begin{cases} \begin{bmatrix} p_z(p_y - b(v)) / (p_x p_y - a(u)b(v)) \\ p_z(p_x - a(u)) / (p_x p_y - a(u)b(v)) \end{bmatrix}, & u \neq c_u, v \neq c_v, \\ \begin{bmatrix} 0 \\ p_z / b(v) \end{bmatrix}, & u = c_u, v \neq c_v, \\ \begin{bmatrix} p_z / a(u) \\ 0 \end{bmatrix}, & u \neq c_u, v = c_v, \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & u = c_u, v = c_v. \end{cases} \quad (26)$$

Мы построили функцию, которая каждому пикселю $g = (u, v)$, заданному в F_{im} , ставит в соответствие пространственную точку в координатах F_c

согласно (25), (26):

$$\xi(\mathbf{g}) = \begin{bmatrix} x_c(u, \nu) \\ y_c(u, \nu) \\ z_c(u, \nu) \end{bmatrix}. \quad (27)$$

Теперь следует уточнить область определения $\xi(\mathbf{g})$, в которой пиксели \mathbf{g} изображают или могут изображать точки плоскости Π (поверхность дороги периодически перекрывается от камеры автомобилями и другими объектами). Обозначим эту область Q и фиксируем ее как четырехугольник на изображении дороги, оригинал которого в Π должен быть близким к прямоугольнику. Выберем четыре угла (g_1, g_2, g_3, g_4) четырехугольника Q (рис. 5). Не составляет труда построить систему из четырех неравенств, которым удовлетворяют пиксели $\mathbf{g} \in Q$. Система F_{im} определяет вещественные координаты, но изображение содержит всего $N \times M$ пикселей. Q содержит часть этого множества – 241072 пикселя. Для $\mathbf{g} \in Q$ можно вычислить соответствующие значения $\xi(\mathbf{g})$. Π – плоскость, поэтому удобно перейти в координаты F_p , привязанные к ней. У всех пространственных точек $(x_p, y_p, z_p) \in \Pi$ в координатах F_p $z_p = 0$. Если точка лежит на объекте, который выше Π (наблюдение точек ниже Π в Q с камеры маловероятно), зачастую удобнее оценивать именно его высоту над поверхностью дороги. Дорожное движение, изображенное в Q , в пространстве можно рассматривать как все, что возвышается над фоном, который лежит на плоскости Π .

Построим F_p . Нормаль к плоскости Π в координатах F_c уже получена в (21), что дает ось Z_p :

$$\mathbf{n}_p = \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix}, \quad \mathbf{e}_{zp} = \mathbf{n}_p / |\mathbf{n}_p|. \quad (28)$$

Остается выбрать в плоскости Π направление еще одной оси, третья определяется автоматически. Пусть ось Y направлена как прообраз вектора $\mathbf{g}_4 - \mathbf{g}_2$ (т.е. вдоль направления движения в области Q), тогда:

$$\mathbf{e}_{yp} = (\xi(\mathbf{g}_4) - \xi(\mathbf{g}_2)) / |\xi(\mathbf{g}_4) - \xi(\mathbf{g}_2)|, \quad \mathbf{e}_{xp} = \mathbf{e}_{yp} \times \mathbf{e}_{zp}. \quad (29)$$

В качестве начала координат выберем

$$o = \xi(\mathbf{g}_2). \quad (30)$$

Матрицы вращения из F_p в F_c и обратно выглядят так:

$$\mathbf{R}_{p2c} = \begin{bmatrix} \mathbf{e}_{xp} & \mathbf{e}_{yp} & \mathbf{e}_{zp} \end{bmatrix} = \begin{bmatrix} e_x^{xp} & e_x^{yp} & e_x^{zp} \\ e_y^{xp} & e_y^{yp} & e_y^{zp} \\ e_z^{xp} & e_z^{yp} & e_z^{zp} \end{bmatrix}, \quad \mathbf{R}_{c2p} = \mathbf{R}_{p2c}^T. \quad (31)$$

Для преобразования координат точки P_c , заданной в F_c в координаты P_p

для системы F_p , используем выражение:

$$P_p = R_{c2p}(P_c - o). \quad (32)$$

Используя (32), (25) – (27), можно перевести координаты всех пикселей Q из F_{im} в метрические координаты F_p . Полученные точки будут лежать в Π .

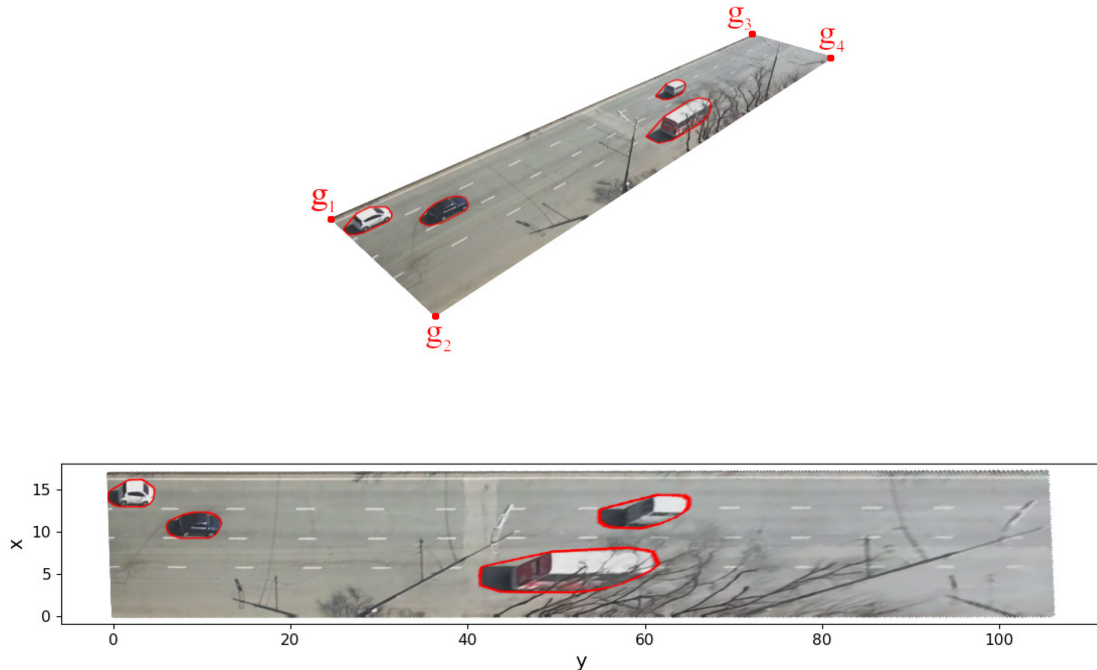


Рис. 5. Визуализация точек в плоскости Π , полученных из пикселей Q с помощью функции ξ (цветом пикселей). Учитывая геометрию образца, координаты x_p выведены по горизонтали, y_p – по вертикали.

Если вывести в этих координатах точки с соответствующим цветом пикселей, получим нижнее изображение рис. 5. Для демонстрации выбрана сцена с различными автомобилями в центре зоны Q . Для пикселей, которые в оригинале лежат на поверхности дороги, мы получили "вид сверху". Для объектов, которые расположены в пространстве выше плоскости Π , изображение искажено. Стоит обратить внимание на цифры на осях координат (это метры). В сравнении с оригиналом (верхнее изображение рис. 5) хорошо заметно, как перспективное искажение меняет восприятие расстояния. Ширина и длина области хорошо соответствуют измерениям, выполненным дальномером и оценкам с online-карт. То, что горизонтальные линии на изображении почти параллельны, говорит о качестве использованных оценок f_u, f_v . Если, например, положить $f_u = f_v = 1300$, интервал значений x_p справа станет заметно уже, чем в начале координат. Таким образом мы выполняем перекрестную проверку используемых значений параметров камеры.

Практически все участники дорожного движения контактируют с поверхностью дороги. Прямую, соединяющую видимые точки контакта колес-

ных пар автомобиля с дорогой, можно считать лежащей на плоскости Π . С подобными интервалами можно выполнять различные измерения движения в серии изображений [1]. Высоту колесных пар (и других частей автомобилей) можно оценивать и компенсировать.

ЛИТЕРАТУРА

1. Зацеровный А.В., Нурминский Е.А. Нейросетевой анализ транспортных потоков городских агломераций на основе данных публичных камер видеонаблюдения // Компьютерные исследования и моделирование. – 2021. – Т.13, №2. – С. 305-318.
2. Camera Calibration and 3D Reconstruction. OpenCV documentation main modules URL: https://docs.opencv.org/4.5.5/d9/d0c/group__calib3d.html (дата обращения 18.05.2022).
3. Camera Calibration. OpenCV tutorials, python URL: http://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html (дата обращения 18.05.2022).
4. Perspective-n-Point (PnP) pose computation. OpenCV documentation URL: https://docs.opencv.org/4.x/d5/d1f/calib3d_solvePnP.html (дата обращения 18.05.2022).
5. Yandex maps, URL: <https://maps.yandex.ru> (дата обращения 18.05.2022).
6. Transformations between ECEF and ENU coordinates. ESA naupedia URL: https://gssc.esa.int/navipedia/index.php/Transformations_between_ECEF_and_ENU_coordinates (дата обращения 18.05.2022).
7. Szeliski R. Computer Vision: Algorithms and Applications // Springer, 2ed, 2022.
8. Peng S., Sturm P. Calibration Wizard: A Guidance System for Camera Calibration Based on Modelling Geometric and Corner Uncertainty // arXiv: 811.03264v2 [cs.CV] 3 Sep 2019.
9. Marchand E., Uchiyama H., Spindler F. Pose estimation for augmented reality: a hands-on survey // IEEE Transactions on Visualization and Computer Graphics. – 2016. – Vol. 22, Iss. 12. DOI: 10.1109/TVCG.2015.2513408.

E-mail:

Зацеровный Александр Владимирович – avz@poi.dvo.ru;

Нурминский Евгений Алексеевич – nurmiskiy.ea@dvfu.ru.