

Обработка и анализ изображений и сигналов

Решение вариационной задачи точка-плоскость ICP на основе комбинирования визуальных и семантических характеристик трехмерной сцены*

А.В. Вохминцев^{1,II}

^I Челябинский государственный университет, г. Челябинск, Россия

^{II} Югорский государственный университет, г. Ханты-Мансийск, Россия

Аннотация. В работе предложен новый точный комбинированный метод решения вариационной задачи точка-плоскость в замкнутой форме для ортогональных преобразований, который используется для построения трехмерной модели окружающей среды по набору изображений и данным карты глубины, полученным с датчиков, расположенных на мобильных платформах. Проведено сравнение предложенного метода с методом Хорна для метрики точка-точка. Результаты компьютерного моделирования показали, что предложенный метод лучше известных методов регистрации, как по точности, так и с точки зрения вычислительной сложности в неконтролируемых условиях.

Ключевые слова: задача регистрации данных, метрика точка-плоскость, локализация, ортогональное преобразование, двумерные дескрипторы, итеративный алгоритм ближайших точек.

DOI: 10.14357/20790279200101

Введение

В общем случае задачу SLAM можно описать как повторяющуюся последовательность шагов: сканирование окружающего пространства мобильной платформой; определение смещения платформы на основе сравнения текущего кадра с предыдущим; выделение на текущем кадре особенностей (меток); сопоставление меток текущего кадра с метками полученными за всю историю наблюдений; обновление на основе этой информации положения платформы за всю историю наблюдений; проверка на замыкание циклов; выравнивание общей карты (отталкиваясь от положения меток и платформы за всю историю наблюдений) [1,2]. Данная работа направлена на разработку ново-

го точного метода регистрации мультисенсорных данных в неизвестной динамической окружающей среде (шаг 4). Важным аспектом, обуславливающим современное состояние задачи трехмерного моделирования окружающей среды, является стремительное развитие различных типов сенсоров. Все большее распространение получают сенсоры глубины и лазерные сканеры, используемые для получения информации о форме трехмерного объекта. Kinect-камера широко используется для получения карты глубины сцены.

В последнее десятилетие было предложено большое количество интересных методов и алгоритмов для решения задачи одновременной навигации и составления карты, однако все известные методы навигации в неизвестном пространстве основаны на использовании визуальной (видимый или инфракрасный диапазон) или одометрической

* Исследование выполнено при финансовой поддержке РФФИ и Ямало-Ненецкого автономного округа в рамках научного проекта №19-47-890001.

информации, а также информации о глубине и их комбинации и не учитывают информацию о семантических свойствах исследуемого неизвестного пространства [3, 4]. Известно, что при построении трехмерных моделей большое распространение получили подходы, основанные на семантической маркировке исследуемого пространства [5]. Данные семантические маркеры могут быть использованы при решении задачи регистрации данных, определения динамического положения платформы в относительной системе координат, а также при решении задачи планирования траектории движения платформы. В этом контексте одной из главных задач является семантическая сегментация и картографирование, а также распознавание ключевых объектов на сцене. В данной работе используется детектор объектов (YOLO) [6] и семантический подход к сегментации, основанный на 3D-запутанных лесах (3DEF) для семантической маркировки пространства [7].

В большинстве случаев, для построения трехмерных моделей по картам глубины используется итеративный алгоритм ближайших точек (Iterative Closest Point, ICP) [8]. Главный этап алгоритма регистрации ICP связан с поиском соответствующего геометрического преобразования (ортогонального или аффинного), которое наилучшим образом совмещает два облака точек в разных RGB-D кадрах для выбранной метрики (вариационная задача алгоритма). Наиболее распространенными вариантами регистрации ICP являются алгоритмы точка-точка (point-to-point) и точка-плоскость (point-to-plane). Разница между этими двумя методами заключается в том, что в рамках подхода на основе точка-плоскость минимизируются расстояния между точками текущего кадра и касательными прямыми к точкам опорного кадра, а в рамках подхода точка-точка используются расстояния между соответствующими точками. Точность реконструкции трехмерной сцены существенно зависит от выбора метрики для оценки геометрического преобразования и метода решения вариационной задачи [9, 10]. Результат применения итерационных методов для решения задачи минимизации выбранного функционала зависит от правильности выбора начального приближения параметров геометрического преобразования: итерационный процесс может сходиться медленно, сходиться к локальному оптимуму или не сходиться. Использование решений вариационной задачи в замкнутой форме позволяет избежать этих проблем [11]. Выбор класса геометрических преобразований также оказывает значительное влияние на результат

реконструкции трехмерной сцены [12, 13]. Для класса ортогональных преобразований решение задачи точка-точка в замкнутой форме получено с помощью кватернионов [14] или с помощью ортогональных матриц [15].

Известно, что метрика точка-плоскость [16] превосходит метрику точка-точка с точки зрения точности и скорости сходимости. Если истинное геометрическое преобразование, связывающее два облака точек, является ортогональным, то применение вариационной задачи для аффинных преобразований даст правильный результат только в том случае, если соответствие между точками двух облаков близко к идеальному [17]. На практике, соответствие между точками двух облаков в большинстве случаев не является идеальным, например, все точки первого облака могут соответствовать одной точке второго облака или небольшому локальному подмножеству точек второго облака. Отметим также, что решение вариационной задачи минимизации функционала в замкнутой форме как условной вариационной задачи с ограничением на принадлежность решения классу ортогональных матриц приводит к необходимости работать с существенно нелинейным подмножеством ортогональных матриц многообразия матриц линейных преобразований [18, 19]. Поэтому рассматриваемая вариационная задача для ортогональных преобразований как математическая проблема значительно сложнее соответствующей задачи для аффинных преобразований.

Для компьютерного моделирования в работе использовалась эталонная база данных, которая содержит наборы кадров сцены в виде данных о цвете и глубине, полученные с использованием камеры Kinect. Основные недостатки итерационных методов регистрации данных связаны с ограничением области сходимости и большой вычислительной сложностью, также результат решения вариационной задачи зависит от правильности выбора начального приближения [20, 21]. Для преодоления данного недостатка в работе предлагается использовать визуально связанные характеристики RGB-D кадра (особые точки), которые позволяют совмещать кадры без требования начальной инициализации. Хорн предложил решение условной вариационной задачи для метрики точка-точка в замкнутой форме для аффинных преобразований, в данной работе предложен новый метод решения вариационной задачи точка-плоскость в замкнутой форме для ортогональных преобразований на основе комбинации визуальных характеристик окружающей среды (данных о глубине и цвете в кадре) и семантических характеристик окружаю-

щей среды. Метод основан на проецировании элемента многообразия матриц линейных преобразований на подмногообразии ортогональных матриц. Предложенный метод реконструкции трехмерной модели сцены комбинирует визуальные и семантические характеристики окружающей среды, для получения которых используются детектор объектов на основе гистограмм ориентированных градиентов [22], итеративный алгоритм ближайших точек [15] и метод семантической сегментации на основе 3D-запутанных лесов [7] соответственно.

1. Общая постановка задачи регистрации. Вариационная задачи алгоритма ICP.

Пусть $X = \{x_1, \dots, x_n\}$ исходное облако точек и $Y = \{y, \dots, y_m\}$ целевое облако точек в $\mathbb{R}^3$. Предположим, что отношение между точками в облаке X и облаке Y такое, что для каждой точки x_i существует точка y_i . Тогда можно рассматривать вариационную задачу ICP как произвольное преобразование ригидных объектов из X в Y :

$$Rx_i + T, \quad (1)$$

где R – матрица поворота, T – вектор переноса, $i = 1, \dots, n$.

Входной информацией для вариационной задачи ICP являются: исходные облака точек X и Y ; первичная оценка для преобразования (начальная инициализация), критерии для останковки итераций. Выходной информацией является такое преобразование R и T , которое сводит к минимуму расстояние между точками из двух необработанных кадров.

Задача регистрации трехмерных данных на основе алгоритма ICP состоит из следующих шагов:

1. Формирование разреженных подмножеств точек из двух плотных трехмерных облаков точек.
2. Определение соответствующих точек в каждом из разреженных подмножеств.
3. Определение весовых коэффициентов для каждой полученной пары в RGB-D кадрах.
4. Отбрасывание некоторых пар в облаках точек с целью исключения выбросов.
5. Выбор метрики ошибки для пар точек.
6. Проецирование элемента многообразия матриц линейных преобразований на подмногообразии ортогональных матриц.
7. Решение вариационной задачи на основе минимизации функции ошибки.

В работе [18] для класса аффинных преобразований было получено следующее решение вариационной задачи точка-точка ICP:

$$E(R, T) = \argmin_{\tilde{M}} \left[\alpha \frac{1}{W} \left(\frac{1}{|A_f|} \sum_i w_i |\tilde{M} F_x^i - F_y^i|^2 \right) + (1 - \alpha) \frac{1}{W} \left(\frac{1}{|A_d|} \sum_j w_j |(\tilde{M} x_j - y_j)|^2 \right) \right], \quad (2)$$

где F_x^i, F_y^i – визуально связанные характеристики сцены (соответствующие точки на изображениях); A_f – подмножество содержит связи между соответствующими точками F_x^i и F_y^i в изображениях в двух кадрах данных; A_d – содержит связи между соответствующими точками x_j и y_j в трехмерных облаках точек в двух кадрах данных; w_j и w_i – весовые характеристики, связанные с семантической маркировкой пространства; $\tilde{M}$ – матрица преобразований, содержащая компоненты поворота R и параллельного переноса T .

2. Точное решение комбинированной вариационной задачи ICP для ортогональных преобразований

Известно, что аффинное преобразование является функцией от двенадцати переменных, тогда можно решить вариационную задачу точка-плоскость алгоритма ICP для произвольного аффинного преобразования. Обозначим через $S(Y)$ поверхность, построенную из облака точек Y , через $T(y^i)$ обозначим касательную плоскость к $S(Y)$ в точке y^i . Пусть x^i есть точка в облаке точек X , тогда n_i – есть унитарная нормаль к y^i :

$$x^i = (x_1^i \ x_2^i \ x_3^i \ 1), \quad y^i = (y_1^i \ y_2^i \ y_3^i \ 1),$$

$$F_x^i = (x f_1^i \ x f_2^i \ x f_3^i \ 1),$$

$$F_y^i = (y f_1^i \ y f_2^i \ y f_3^i \ 1), \quad n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 1 \end{pmatrix}. \quad (3)$$

Представим графическую интерпретацию метрики точка-плоскость между двумя облаками

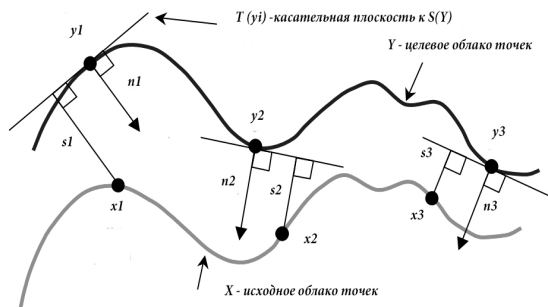


Рис. 1. Метрика точка-плоскость между двумя облаками точек

точек (Рис. 1) [16] в двумерном пространстве для простоты восприятия.

Тогда можно представить вариационную задачу точка-плоскость в следующем виде:

$$J(RV, RD) = \alpha \frac{1}{W} \frac{1}{|A_f|} \sum_{i=1}^n w_i (\langle RV * F_x^i - F_y^i n^i \rangle)^2 + (1 - \alpha) \frac{1}{W} \frac{1}{|A_d|} \sum_{i=1}^n w_i (\langle RD * x^i - y^i n^i \rangle)^2, \quad (4)$$

где $\langle \cdot, \cdot \rangle$ есть скалярное произведение, RV – матрица аффинного преобразования для визуально связанных характеристик сцены, RD – матрица аффинного преобразования для данных о глубине сцены:

$$RD = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, RV = \begin{pmatrix} f_{11} & f_{12} & f_{13} & b_1 \\ f_{21} & f_{22} & f_{23} & b_2 \\ f_{31} & f_{32} & f_{33} & b_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5)$$

Обозначим $\eta_1 = \alpha \frac{1}{W} \frac{1}{|A_f|}$ и $\eta_2 = (1 - \alpha) \frac{1}{W} \frac{1}{|A_d|}$.

Представим функцию $J(RV, RD)$ в следующем виде:

$$\begin{aligned} J(RV, RD) &= \eta_1 \sum_{i=1}^n w_i (\langle RV F_x^i - F_y^i n^i \rangle)^2 + \eta_2 \sum_{i=1}^n w_i (\langle RD x^i - y^i n^i \rangle)^2 \\ &= \eta_1 \sum_{i=1}^n w_i (\langle RV F_x^i, n^i \rangle - \langle F_y^i, n^i \rangle)^2 \\ &+ \eta_2 \sum_{i=1}^n w_i (\langle RD x^i, n^i \rangle - \langle y^i, n^i \rangle)^2 \\ &= \eta_1 \sum_{i=1}^n (\langle RV F_x^i, n^i \rangle)^2 - 2 \langle RV F_x^i, n^i \rangle \langle F_y^i, n^i \rangle + (\langle F_y^i, n^i \rangle)^2 \\ &+ \eta_2 \sum_{i=1}^n (\langle RD x^i, n^i \rangle)^2 - 2 \langle RD x^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2. \end{aligned} \quad (6)$$

Выражения $\langle F_y^i, n^i \rangle$ и $\langle y^i, n^i \rangle$ являются постоянными относительно RV и RD соответственно, следовательно вариационная задача (4) может быть представлена в следующем виде:

$$\eta_1 \sum_{i=1}^n (\langle RV F_x^i, n^i \rangle)^2 - 2 \langle RV F_x^i, n^i \rangle \langle F_y^i, n^i \rangle + \eta_2 \sum_{i=1}^n (\langle RD x^i, n^i \rangle)^2 - 2 \langle RD x^i, n^i \rangle \langle y^i, n^i \rangle \rightarrow \min_{RV, RD}. \quad (7)$$

Давайте сделаем некоторые преобразования над скалярными произведениями $\langle RV F_x^i, n^i \rangle$ и $\langle RD x^i, n^i \rangle$ и представим их как:

$$\begin{aligned} \langle RV F_x^i, n^i \rangle &= (f_{11} F_{1x}^i n_1^i + f_{12} F_{2x}^i n_1^i + f_{13} F_{3x}^i n_1^i + b_1 n_1^i) + \\ &+ (f_{21} F_{1x}^i n_2^i + f_{22} F_{2x}^i n_2^i + f_{23} F_{3x}^i n_2^i + b_2 n_2^i) + \\ &+ (f_{31} F_{1x}^i n_3^i + f_{32} F_{2x}^i n_3^i + f_{33} F_{3x}^i n_3^i + b_3 n_3^i) + (0) \end{aligned} \quad (8) \quad \text{и}$$

$$\begin{aligned} \langle RD x^i, n^i \rangle &= (r_{11} x_1^i n_1^i + r_{12} x_2^i n_1^i + r_{13} x_3^i n_1^i + t_1 n_1^i) + \\ &+ (r_{21} x_1^i n_2^i + r_{22} x_2^i n_2^i + r_{23} x_3^i n_2^i + t_2 n_2^i) + \\ &+ (r_{31} x_1^i n_3^i + r_{32} x_2^i n_3^i + r_{33} x_3^i n_3^i + t_3 n_3^i) + (0). \end{aligned} \quad (9)$$

Давайте рассмотрим компоненты градиента $\nabla(RV, RD)$. Для начала давайте рассмотрим компоненты градиента относительно переменных t и b соответственно:

$$\frac{\partial J}{\partial t_j} = \sum_{i=1}^n 2 \langle RD x^i, n^i \rangle n_j^i - 2 n_j^i \langle y^i, n^i \rangle = 0, j = 1, 2, 3 \quad (10)$$

и

$$\frac{\partial J}{\partial b_j} = \sum_{i=1}^n 2 \langle RV F_x^i, n^i \rangle n_j^i - 2 n_j^i \langle F_y^i, n^i \rangle = 0, j = 1, 2, 3. \quad (11)$$

Далее воспользуемся свойствами суммы:

$$\sum_{i=1}^n \langle RD x^i, n^i \rangle n_j^i - n_j^i \langle y^i, n^i \rangle = \sum_{i=1}^n \langle RD x^i, n^i \rangle n_j^i - \sum_{i=1}^n n_j^i \langle y^i, n^i \rangle = 0. \quad (12)$$

Сумма $\sum_{i=1}^n n_j^i \langle y^i, n^i \rangle$ является постоянной относительно RD , тогда обозначим сумму следующим образом $\sum_{i=1}^n n_j^i \langle y^i, n^i \rangle = \alpha_j$. Тогда, с учетом этого определения выражение (12) может быть представлено в следующей форме:

$$\alpha_j = \sum_{i=1}^n \langle RD x^i, n^i \rangle n_j^i. \quad (13)$$

Аналогично сумма $\sum_{i=1}^n n_j^i \langle F_y^i, n^i \rangle$ является постоянной относительно RV , тогда

$$\beta_j = \sum_{i=1}^n \langle RV F_x^i, n^i \rangle n_j^i. \quad (14)$$

Далее обозначим через $(XN)^i$ и $(FXN)^i$ следующие матрицы соответственно:

$$(XN)^i = x^i n^i = \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix} \quad (15)$$

и

$$(FXN)^i = F_x^i n^i = \begin{pmatrix} F_{1x}^i n_1^i & F_{1x}^i n_2^i & F_{1x}^i n_3^i & 0 \\ F_{2x}^i n_1^i & F_{2x}^i n_2^i & F_{2x}^i n_3^i & 0 \\ F_{3x}^i n_1^i & F_{3x}^i n_2^i & F_{3x}^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \quad (16)$$

Затем рассмотрим выражения следующего вида:

$$\langle RD x^i, n^i \rangle = \text{Tr}(RD \cdot (XN)^i) \quad (17)$$

$$\langle RV F_x^i, n^i \rangle = \text{Tr}(RV \cdot (FXN)^i), \quad (18)$$

где Tr – операция след матрицы.

Давайте представим выражения (13) и (14) с учетом выражений (17) и (18) соответственно как:

$$\sum_{i=1}^n n_j^i \text{Tr}(RD \cdot (XN)^i) = \alpha_j, \sum_{i=1}^n n_j^i \text{Tr}(RV \cdot (FXN)^i) = \beta_j. \quad (19)$$

Обозначим через $(n_j XN)^i$ и $(n_j FXN)^i$ следующие матрицы:

$$(n_j XN)^i = n_j^i (XN)^i = n_j^i \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix} \quad (20)$$

и

$$(n_j FXN)^i = n_j^i (FXN)^i = n_j^i \begin{pmatrix} F_{1x}^i n_1^i & F_{1x}^i n_2^i & F_{1x}^i n_3^i & 0 \\ F_{2x}^i n_1^i & F_{2x}^i n_2^i & F_{2x}^i n_3^i & 0 \\ F_{3x}^i n_1^i & F_{3x}^i n_2^i & F_{3x}^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \quad (21)$$

Заметим, что выполняются следующие условия

$$n_j^i \text{Tr}(RD \cdot (XN)^i) = \text{Tr}(RD \cdot (n_j XN)^i) \quad (22)$$

$$n_j^i \text{Tr}(RV \cdot (FXN)^i) = \text{Tr}(RV \cdot (n_j FXN)^i). \quad (23)$$

С учетом условия в выражениях (22) и (23) представим выражение (19) в следующем виде:

$$\sum_{i=1}^n \text{Tr}(RD \cdot (n_j XN)^i) = \alpha_j$$

и

$$\sum_{i=1}^n \text{Tr}(RV \cdot (n_j FXN)^i) = \beta_j. \quad (24)$$

Давайте используем свойство трассировки матрицы, тогда получим в результате некоторых преобразований выражения вида:

$$\begin{aligned} \sum_{i=1}^n \text{Tr}(RD \cdot (n_j XN)^i) &= \text{Tr}(\sum_{i=1}^n RD \cdot (n_j XN)^i) = \\ \text{Tr}(RD (\sum_{i=1}^n (n_j XN)^i)) &= \alpha_j \end{aligned} \quad (25)$$

и

$$\begin{aligned} \sum_{i=1}^n \text{Tr}(RV \cdot (n_j FXN)^i) &= \text{Tr}(\sum_{i=1}^n RV \cdot (n_j FXN)^i) = \\ \text{Tr}(RV (\sum_{i=1}^n (n_j FXN)^i)) &= \beta_j \end{aligned} \quad (26)$$

Давайте обозначим через SV^j и SD^j , $j = 1, 2, 3$ следующие матрицы:

$$\sum_{i=1}^n (n_j XN)^i = SD^j, \sum_{i=1}^n (n_j FXN)^i = SV^j, \quad (27)$$

где

$$SD^j = \begin{pmatrix} s_{11}^{jd} & s_{12}^{jd} & s_{13}^{jd} & 0 \\ s_{21}^{jd} & s_{22}^{jd} & s_{23}^{jd} & 0 \\ s_{31}^{jd} & s_{32}^{jd} & s_{33}^{jd} & 0 \\ s_{41}^{jd} & s_{42}^{jd} & s_{43}^{jd} & 0 \end{pmatrix},$$

$$SV^j = \begin{pmatrix} s_{11}^{jv} & s_{12}^{jv} & s_{13}^{jv} & 0 \\ s_{21}^{jv} & s_{22}^{jv} & s_{23}^{jv} & 0 \\ s_{31}^{jv} & s_{32}^{jv} & s_{33}^{jv} & 0 \\ s_{41}^{jv} & s_{42}^{jv} & s_{43}^{jv} & 0 \end{pmatrix}. \quad (28)$$

Тогда выражения (25) и (26) могут быть представлены в следующем виде соответственно:

$$\text{Tr}(RD \cdot SD^j) = \alpha_j \text{ и } \text{Tr}(RV \cdot SV^j) = \beta_j. \quad (29)$$

Найдем частые производные функционала $J(RD)$ относительно r_{ij} , f_{ij} , $i, j = 1, 2, 3, k = 1, \dots, n$:

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^n 2 \langle RD x^k, n^k \rangle x_j^k n_i^k - 2 x_j^k n_i^k \langle y^k, n^k \rangle = 0 \quad (30)$$

и

$$\frac{\partial J}{\partial f_{ij}} = \sum_{k=1}^n 2 \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - 2 F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = 0. \quad (31)$$

Тогда

$$\begin{aligned} \sum_{k=1}^n \langle RD x^k, n^k \rangle x_j^k n_i^k - x_j^k n_i^k \langle y^k, n^k \rangle \\ = \sum_{k=1}^n \langle RD x^k, n^k \rangle x_j^k n_i^k - \sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = 0, \end{aligned} \quad (32)$$

$$\begin{aligned} \sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - F_{jx}^k n_i^k \langle F_y^k, n^k \rangle \\ = \sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k - \sum_{k=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = 0. \end{aligned} \quad (33)$$

Частная сумма $\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle$ есть константа относительно RD , а частная сумма $\sum_{i=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle$ есть константа относительно RV , тогда:

$$\sum_{k=1}^n x_j^k n_i^k \langle y^k, n^k \rangle = \alpha_{ij}$$

и

$$\sum_{k=1}^n F_{jx}^k n_i^k \langle F_y^k, n^k \rangle = \beta_{ij}. \quad (34)$$

Тогда выражение (33) может быть представлено в виде:

$$\sum_{k=1}^n \langle RDx^k, n^k \rangle x_j^k n_i^k = \alpha_{ij}, \quad (35)$$

$$\sum_{k=1}^n \langle RV F_x^k, n^k \rangle F_{jx}^k n_i^k = \beta_{ij}. \quad (36)$$

Принимая во внимание выражение (17), выражение (35) может быть представлено как:

$$\sum_{k=1}^n x_j^k n_i^k \text{Tr}(RD \cdot (XN)^k) = \alpha_{ij} \quad (37)$$

и выражение (18) с учетом (36) может быть представлено как

$$\sum_{k=1}^n F_{jx}^k n_i^k \text{Tr}(RV \cdot (FXN)^k) = \beta_{ij}. \quad (38)$$

Обозначим через $x_j^k n_i^k XN^k$ и $F_{jx}^k n_i^k FXN^k$ следующие матрицы

$$x_j^k n_i^k XN^k = x_j^k n_i^k \begin{pmatrix} x_1^k n_1^k & x_1^k n_2^k & x_1^k n_3^k & 0 \\ x_2^k n_1^k & x_2^k n_2^k & x_2^k n_3^k & 0 \\ x_3^k n_1^k & x_3^k n_2^k & x_3^k n_3^k & 0 \\ n_1^k & n_2^k & n_3^k & 0 \end{pmatrix} \quad (39)$$

и

$$F_{jx}^k n_i^k FXN^k = F_{jx}^k n_i^k \begin{pmatrix} F_1^k n_1^k & F_1^k n_2^k & F_1^k n_3^k & 0 \\ F_2^k n_1^k & F_2^k n_2^k & F_2^k n_3^k & 0 \\ F_3^k n_1^k & F_3^k n_2^k & F_3^k n_3^k & 0 \\ n_1^k & n_2^k & n_3^k & 0 \end{pmatrix}. \quad (40)$$

Тогда выражения (37) и (38) принимают соответственно вид:

$$\sum_{k=1}^n \text{Tr}(RD \cdot x_j^k n_i^k XN^k) = \alpha_{ij} \quad (41)$$

и

$$\sum_{k=1}^n \text{Tr}(RV \cdot F_{jx}^k n_i^k FXN^k) = \beta_{ij}. \quad (42)$$

Далее обозначим SV^{ij} и SD^{ij} , $i, j = 1, 2, 3$, следующие матрицы, тогда:

$$\sum_{k=1}^n x_j^k n_i^k XN^k = SD^{ij}$$

и

$$\sum_{k=1}^n F_{jx}^k n_i^k FXN^k = SV^{ij}. \quad (43)$$

3. Решение вариационной задачи точка-плоскость метода ICP для ортогональных преобразований

На основе сказанного в разделе 2 может быть получено множество из 24 линейных уравнений относительно переменных $r_{ij}, f_{ij}, t_k, b_k, i, j, k = 1, \dots, 3$. Таким образом, можно представить решение комбинированной вариационной задачи для ортогональных преобразований в матричной форме:

$$\begin{cases} \text{Tr}(RD \cdot SD^j) = \alpha_j \\ \text{Tr}(RV \cdot SV^j) = \beta_j \\ \sum_{k=1}^n \text{Tr}(RD \cdot SD^{ij}) = \alpha_{ij} \Rightarrow Z \cdot \Omega = \theta, \\ \sum_{k=1}^n \text{Tr}(RV \cdot SV^{ij}) = \beta_{ij} \end{cases} \quad (44)$$

где матрица Z имеет размер 18×24 и состоит из элементов SV^{ij} и SD^{ij} ; θ есть матрица из 24 столбцов $\alpha_j, \beta_j, \alpha_{ij}, \beta_{ij}$ и Ω есть вектор следующего вида:

$$\Omega = (r_{11}, r_{12}, r_{13}, t_1, r_{21}, r_{22}, r_{23}, t_2, r_{31}, r_{32}, r_{33}, t_3, f_{11}, f_{12}, f_{13}, b_1, f_{21}, f_{22}, f_{23}, b_2, f_{31}, f_{32}, f_{33}, b_3). \quad (45)$$

В общем случае, $RV \neq RD$, в данной работе будем искать совместное решение вариационной задачи для ортогональных преобразований для частного случая, когда $RV = RD = RT$. Пусть T_{km} – начальная оценка для ICP с использованием кинематической модели движения камеры, $k_{\max}$, ε – пороги алгоритма ICP по числу шагов и по величине ошибки E соответственно, RT^* – лучшее преобразование на i -ом шаге метода, δ – порог для точек инлайнеров.

Общая схема комбинированного метода регистрации (решение вариационной задачи) для ортогональных преобразований может быть представлена как:

Шаг 1. Вычисление гистограмм направленных градиентов на изображениях.

Шаг 2. Сопоставление между особыми точками F_x^i и F_y^i для выбранных подмножеств. Решение вариационной задачи регистрации данных для визуально связанных характеристик изображения. Получим RT^* , A_r .

Шаг 3. Проверка: Если $|A_r| < \delta$, то $RT^* = T_{km}$ и $A_r = \emptyset$. Положим $i=1$.

Шаг 4. Определение соответствующих точек A_d в исходном и целевом облаке точек с использованием метода ближайших соседей. Определение весовых коэффициентов для каждой пары A_d .

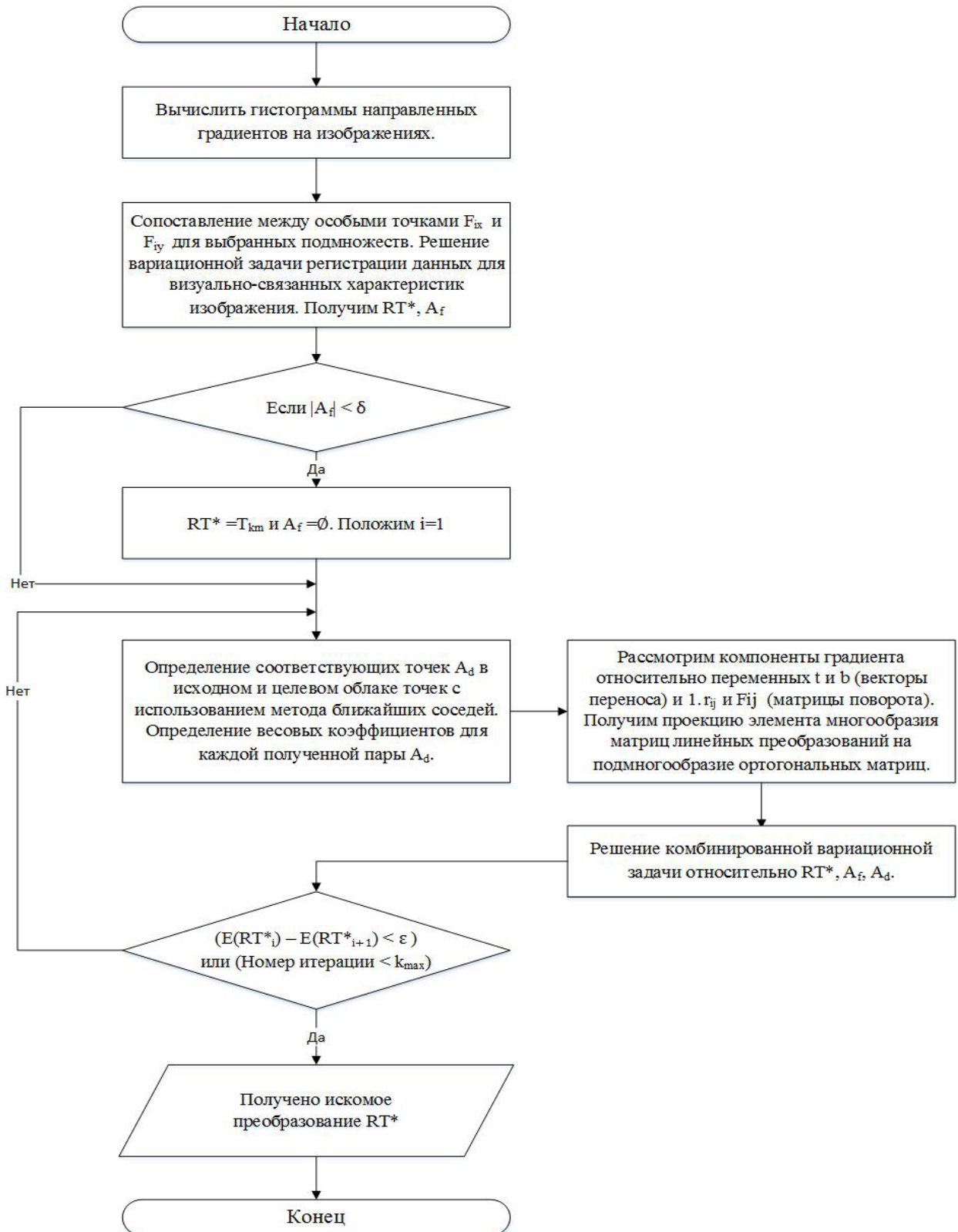


Рис. 2. Схема алгоритма для решения комбинированной задачи регистрации данных для ортогональных преобразований

Шаг 5. Рассмотрим компоненты градиента относительно переменных t и b : векторы переноса, формулы (10) и (11), и относительно переменных r_{ij}, f_{ij} : матрицы поворота, формулы (30) и (31), соответственно. Получим подмножество ортогональных матриц $n_i^j(XN)^i$, $n_j^i(FXN)^i$ и $x_j^k n_i^k XN^k$, $F_{jx}^k n_i^k FXN^k$ (см. формулы (20), (21), (39), (40) соответственно).

Шаг 6. Решение комбинированной вариационной задачи относительно RT^* , A_f , A_d в соответствии с системой уравнений в формуле (44).

Шаг 7. Проверка: $(E(RT_i^*) - E(RT_{i+1}^*)) \leq \epsilon$ или (Номер итерации $< k_{\max}$). Если условие не верно, то возврат на Шаг 4 и $i=i+1$, иначе получено искомое преобразование RT^* .

Шаг 5 обладает большой вычислительной сложностью, поэтому в общем случае решение комбинированной задачи регистрации для аффинных преобразований является менее затратной процедурой, чем решение аналогичной комбинированной задачи регистрации для ортогональных преобразований. Общая схема алгоритма для решения комбинированной задачи регистрации данных для ортогональных преобразований представлена на Рис 2.

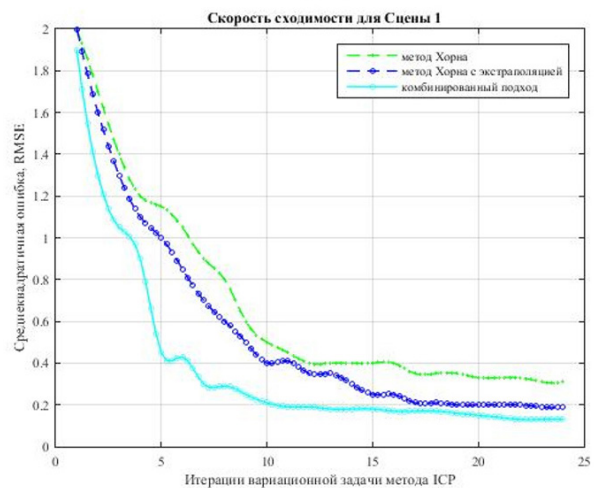
4. Компьютерное моделирование

В данном разделе представлены и обсуждены результаты компьютерного моделирования. Для проведения компьютерного моделирования было выбрано четыре набора данных из базы данных NYU Depth Dataset V2 [23], содержащие фрагменты крупномасштабных сцен помещений: Classrooms, Living Rooms (1/4), Offices (1/2), Offices (2/2). Компьютерное моделирование проводилось на персональном компьютере Intel Core i7 с многоядерным графическим процессором. Экспериментальные данные были получены с использованием камеры Kinect в качестве RGB-D сенсора. В работе проведен сравнительный анализ скорости сходимости предложенного метода регистрации данных для ортогональных преобразований и известных методов регистрации на основе итеративного алгоритма ближайших точек в зависимости от выбора метрики ошибки пар точек: точка-точка (метод Хорна) [14], точка-точка с экстраполяцией (метод Хорна с экстраполяцией) [15] в терминах среднеквадратичной ошибки. Для проведения компьютерного моделирования было выбрано два набора данных из базы данных NYU Depth Dataset: в контролируемых условиях (Offices (1/2), Сцена 1) и в неконтролируемых условиях (Offices (1/2), Сцена 2).

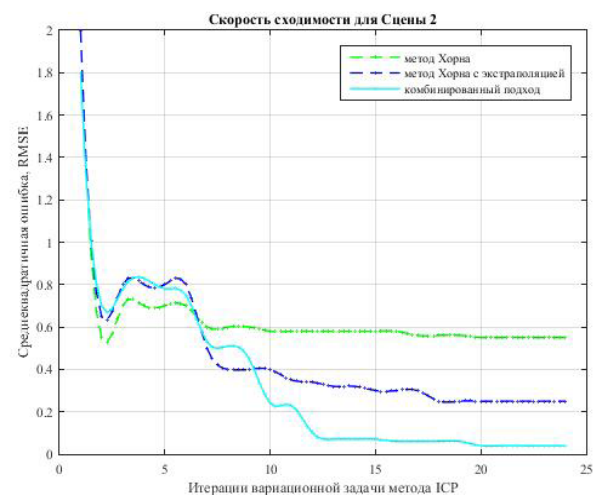
Каждый из четырех наборов данных из тестовой базы содержит файлы дампа, изображение в формате ppm, данные о глубине в формате pgm. Данные в RGB-D кадрах в эталонной базе данных получены из разных положений камеры Kinect на сцене. В результате серии тестов установлены зависимости скорости сходимости методов от выбора ошибки метрики и условий проведения экспериментов (Рис. 3).

Был проведен сравнительный анализ зависимости скорости сходимости методов регистрации для ортогональных преобразований от выбора типа дескриптора и условий проведения экспериментов (Рис. 4).

Было установлено, что в контролируемых условиях тип используемого дескриптора имеет



(а)



(б)

Рис. 3. Сравнение скорости сходимости в зависимости от ошибки метрики и условий наблюдения: (а) изменение скорости сходимости в контролируемых условиях; (б) изменение скорости сходимости в неконтролируемых условиях

ограниченное влияние на сходимость метода регистрации данных (Рис. 4, а), в неконтролируемых условиях использование дескриптора HOGs позволяет получить лучшую сходимость и точность в сравнении с другими дескрипторными методами: предложенный метод регистрации сходится уже после 10 итерации, тогда как при использовании дескриптора ORB предложенный метод регистрации сходится только после 14 итерации (Рис. 4, б).

Для получения указанного результата в работе проведено исследование точности трехмерной реконструкции от количества и пространственного распределения особых точек в кадре. Было установлено, что точность метода регистрации зависит от количества особых точек в кадре нелинейным

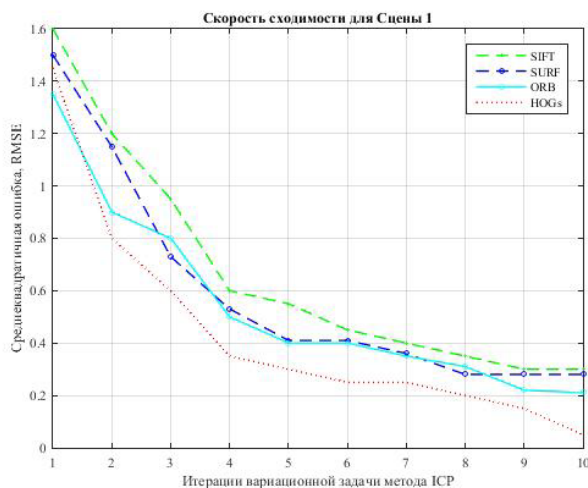
образом – в виде функции с одним ярко выраженным пиком для всех типов дескрипторов. Было установлено, что приближенные методы решения позволяют получить достаточно точные результаты при реконструкции трехмерных сцен окружающего пространства. Таким образом, можно сделать вывод, что применение визуально связанных характеристик для решения вариационной задачи алгоритма ICP позволяет получать точное решение вариационной задачи регистрации для заданной трехмерной сцены. Предложенные методы сопоставления и регистрации RGB-D кадров являются локальными, поэтому возможна их параллельная реализация с использованием персональных компьютеров с многоядерным графическим процессором, имеющим интегрированную параллельную архитектуру. Для обеспечения режима реального времени работы (30 кадров в секунду) на персональном компьютере с многоядерным графическим процессором алгоритмы используют SIMD технологию.

Заключение

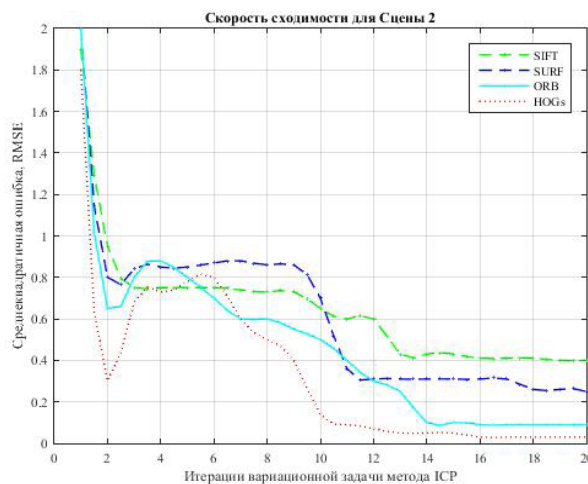
В работе проведен сравнительный анализ известных методов регистрации на основе вариантов метода ICP, было установлено, что сходимость существующих методов может быть значительно улучшена. В современных алгоритмах часто вместо соответствия точка к точке используется соответствие точка к плоскости. При таком подходе ошибка, которую минимизирует метод, является расстоянием между точками текущего скана и касательными прямыми к точкам опорного скана. Однако подход точка-плоскость накладывает более жесткие ограничения на структуру реконструируемого пространства, такой вариант алгоритма ICP имеет плохую сходимость на сценах с малым количеством геометрических ограничений. В работе предложен новый метод решения вариационной задачи точка-плоскость для класса ортогональных преобразований без требования начальной инициализации. Было проведено сравнение предлагаемого метода регистрации данных с методом Хорна для метрики точка-точка с экстраполяцией и без экстраполяции. В результате компьютерного моделирования было установлено, что применение визуально связанных характеристик для решения вариационной задачи алгоритма ICP позволяет улучшить точность и сходимость метода в контролируемых и неконтролируемых условиях.

Литература

1. Davison A. J., Reid I. D., Molton N. D., Stasse O. MonoSLAM: Real-Time single camera SLAM //



(а)



(б)

Рис. 4. Сравнение скорости сходимости в зависимости от выбора типа дескриптора и условий наблюдения: (а) изменение скорости сходимости в контролируемых условиях; (б) изменение скорости сходимости в неконтролируемых условиях

- Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007. № 29 (6), p. 1052-1067.
2. *Hertzberg C., Wagner R., Birbach O.* Experiences in building a visual slam system from open source components // IEEE International Conference on Robotics and Automation Proceedings. 2011, p. 2644-2651.
 3. *Vokhmintsev A., Yakovlev K.* A Real-time Algorithm for Mobile Robot Mapping Based on Rotation-invariant Descriptors and ICP // The 5th International Conference on Analysis of Images, Social Networks and Texts / Eds. D. I. Ignatov, M. Yu. Khachay, V. G. Labunets, et al. – Communications in Computer and Information Science. – Springer. 2016, 661, p. 357–369.
 4. *Henry P., Krainin M., Herbst E., Ren X., Fox D.* RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments // 12-th International Symposium on Experimental Robotics Proceedings. 2014, 79, p. 477-491.
 5. *Endres F., Hess J., Engelhard N., Sturm J.* An evaluation of the RGB-D SLAM system // IEEE International Conference on Robotics and Automation Proceedings. 2012. P. 1691-1696.
 6. *Redmon J., Farhadi A.* YOLO9000: Better, Faster, Stronger // IEEE Conference on Computer Vision and Pattern Recognition-CVPR Proceedings. 2017.
 7. *Antonello M., Wolf D., Prankl J., Ghidoni S., Menegatti E., Vincze M.* Multi-view 3d entangled forest for semantic segmentation and mapping // IEEE International Conference in Robotics and Automation Proceedings. 2018, p. 1-8.
 8. *Besl P., McKay N.* A method for registration of 3D shapes // Journal of IEEE Transaction on Pattern Analysis and Machine Intelligence. 1992. № 14(2), p. 239-256.
 9. *Tam G.K., Cheng Z.Q., Lai Y.K., Langbein F.C., Liu Y., Marshall D., Martin R.R., Sun X.F., Rosin P.L.* Registration of 3D point clouds and meshes: A survey from rigid to nonrigid // Journal of IEEE Transactions on Visualization and Computer Graphics. 2013. № 19 (7), p. 1199–1217.
 10. *Rusinkiewicz S., Levoy M.* Efficient variants of the Iterative closest point algorithm // 3-rd International Conference on 3D Digital Imaging and Modeling Proceedings. 2001, p. 145–152.
 11. *Khoshelham K.* Closed-form solutions for estimating a rigid motion from plane correspondences extracted from point clouds // ISPRS Journal of Photogrammetry and Remote Sensing. 2016. № 114, p. 78–91.
 12. *Du S., Zheng N., Ying S., Liu J.* Affine iterative closest point algorithm for point set registration // Journal of Pattern Recognition Letters. 2010. № 31(9), p. 791–799.
 13. *Chen Y., Medioni G.* Object modeling by registration of multiple range images // IEEE International Conference on Robotics and Automation Proceedings. 1991, № 10(3), p. 2724-2729.
 14. *Horn B.* Closed-Form Solution of Absolute Orientation Using Unit Quaternions // Journal of the Optical Society of America A. № 4(4), p. 629–642.
 15. *Horn B., Hilden H., Negahdaripour S.* Closed-form Solution of Absolute Orientation Using Orthonormal Matrices // Journal of the Optical Society of America A. 1988, № 5(7), p. 1127–1135.
 16. *Low K.L.* Linear least-squares optimization for point-to-plane ICP surface registration. Department of Computer Science, University of North Carolina at Chapel Hill. Technical Report TTR04-004. Available at: https://www.comp.nus.edu.sg/~lowkl/publications/lowk_point-to-plane_icp_techrep.pdf (accessed February 17, 2004).
 17. *Segal A., Haehnel D., Thrun S.* Generalized-ICP // Science and Systems Proceedings of Robotics. 2009, 2(4), p. 645-653.
 18. *Vokhmintsev A.V., Melnikov A.V., Mironov K.V., Burlutskiy V.V.* Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration // Journal of Doklady Mathematics. 2019. № 99 (1), p. 108–112.
 19. *Vokhmintsev A., Melnikov A., Pachganov S., Burlutskii V.* The New Combined Closed-Solution for 3D Reconstruction of Environment Based on Iterative Closest Point Algorithm // 7 th Scientific Conference on Information Technologies for Intelligent Decision Making Support Proceedings. 2019, p. 23-27.
 20. *Langis C., Greenspan M., Godin G.* The parallel Iterative closest point algorithm // IEEE 3-rd International Conference 3D Digital Imaging and Modeling Proceedings. 2001, p. 195-204.
 21. *Thrun S., Burgard W., Fox D.* A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping // IEEE International Conference on Robotics and Automation Proceedings. 2000, p. 321–328.
 22. *Vokhmintsev A., Sochenkov I., Kuznetsov V., Tikhonkikh D.* Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram // Journal of Doklady Mathematics. 2016. № 93 (1), p. 37-41.
 23. *New York University depth dataset V2.* Available at: https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html (accessed December 28, 2012).

Вохминцев Александр Владиславович. ФГБОУ ВО «Челябинский государственный университет». г. Челябинск, Россия. Заведующий лабораторией. Количество печатных работ: 39. Область научных интересов: цифровая обработка информации, методы сопоставления и регистрации изображений, методы одновременной навигации и составления карты в неизвестном пространстве, распознавание личности по лицу. E-mail: vav@csu.ru

The solution of the variational problem point-plane ICP based on the fusion of visual and semantic characteristics of a three-dimensional scene

A.V. Vokhmintcev^{I,II}

^I Chelyabinsk State University

^{II} Ugra Research Institute of Information Technologies

Abstract. A new accurate combined method for orthogonal transformations is developed for solving the point-plane variational problem in a closed form. This method is used to reconstruct a three-dimensional model of the environment from a set of images and depth map obtained from sensors which located on mobile platforms. The suggested method was compared with the Horn method for the point-to-point metric. The results of computer simulation showed that the proposed method is better than the state-of-art methods of registration both in accuracy and in terms of computational complexity in uncontrolled conditions.

Keywords: *registration task, point-plane metric, localization, orthogonal transformation, two-dimensional descriptors, iterative closest points algorithm*

DOI: 10.14357/20790279200101

References

1. Davison, A. J., Reid, I. D., Molton, N. D., and Stasse, O. 2007. MonoSLAM: Real-Time single camera SLAM. *Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence*. 29 (6): 1052-1067. doi: 10.1109/TPAMI.2007.1049.
2. Hertzberg, C., Wagner, R., and Birbach, O. 2011. Experiences in building a visual slam system from open source components. 2011. *IEEE International Conference on Robotics and Automation Proceedings*. Shanghai, China. 2644-2651.
3. Vokhmintsev, A., and Yakovlev, K. 2016. A Real-time Algorithm for Mobile Robot Mapping Based on Rotation-invariant Descriptors and ICP. *The 5th International Conference on Analysis of Images, Social Networks and Texts / Eds. D. I. Ignatov, M. Yu. Khachay, V. G. Labunets, et al. – Communications in Computer and Information Science*. – Springer. 661: 357–369.
4. Henry, P., Krainin, M., Herbst, E., Ren, X., and Fox, D. 2014. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments. *12th International Symposium on Experimental Robotics Proceedings*. Delhi, India. 79: 477-491.
5. Endres, F., Hess, J., Engelhard, N., and Sturm, J. 2012. An evaluation of the RGB-D SLAM system. *IEEE International Conference on Robotics and Automation Proceedings*. Taipei, Taiwan: 1691-1696.
6. Redmon, J., and Farhadi, A. 2017. YOLO9000: Better, Faster, Stronger. *IEEE Conference on Computer Vision and Pattern Recognition-CVPR Proceedings*. Honolulu, HI, USA.
7. Antonello, M., Wolf, D., Prankl, J., Ghidoni, S., Menegatti, E., and Vincze, M. 2018. Multi-view 3d entangled forest for semantic segmentation and mapping. *IEEE International Conference in Robotics and Automation Proceedings*. Brisbane, QLD, Australia: 1-8.
8. Besl, P., and McKay, N. 1992. A method for registration of 3D shapes. *Journal of IEEE Transaction on Pattern Analysis and Machine Intelligence*. 14(2): 239-256. doi: 10.1109/34.121791.
9. Tam, G.K., Cheng, Z.Q., Lai, Y.K., Langbein, F.C., Liu, Y., Marshall, D., Martin, R.R, Sun, X.F., and Rosin, P.L. 2013. Registration of 3D point clouds and meshes: A survey from rigid to nonrigid. *Journal of IEEE Transactions on Visualization and Computer Graphics*. 19 (7): 1199–1217. doi: 10.1109/TVCG.2012.310
10. Rusinkiewicz, S., and Levoy, M. Efficient variants of the Iterative closest point algorithm. 2001. *3-rd International Conference on 3D Digital Imaging and Modeling Proceedings*. Quebec City, Canada: 145–152.
11. Khoshelham, K. 2016. Closed-form solutions for estimating a rigid motion from plane correspondences extracted from point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*. 114: 78–91. doi: 10.1016/j.

- isprsjprs.2016.01.010.
12. Du, S., Zheng, N., Ying, S., and Liu, J. 2010. Affine iterative closest point algorithm for point set registration. *Journal of Pattern Recognition Letters*. 31(9): 791–799. doi: 10.1016/j.patrec.2010.01.020.
 13. Chen, Y., and Medioni, G. 1991. Object modeling by registration of multiple range images. *IEEE International Conference on Robotics and Automation Proceedings*. Sacramento, USA. 10(3): 2724–2729.
 14. Horn, B. 1987. Closed-Form Solution of Absolute Orientation Using Unit Quaternions. *Journal of the Optical Society of America A*. 4(4): 629–642. doi: 10.1364/josaa.4.000629.
 15. Horn, B., Hilden, H., and Negahdaripour, S. 1988. Closed-form Solution of Absolute Orientation Using Orthonormal Matrices. *Journal of the Optical Society of America A*. 5(7): 1127–1135. doi: 10.1364/JOSAA.5.001127.
 16. Low, K. L. 2004. Linear least-squares optimization for point-to-plane ICP surface registration. Department of Computer Science, University of North Carolina at Chapel Hill. Technical Report TTR04-004. Available at: https://www.comp.nus.edu.sg/~lowkl/publications/lowk_point-to-plane_icp_techrep.pdf (accessed February 17, 2004).
 17. Segal, A., Haehnel, D., and Thrun, S. 2009. Generalized-ICP. *Science and Systems Proceedings of Robotics*. Seattle, USA. 2(4): 645–653.
 18. Vokhmintcev, A.V., Melnikov, A.V., Mironov, K.V., and Burlutskiy, V.V. 2019. Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration. *Journal of Doklady Mathematics*. 99(1): 108–112. doi: 10.1134/S1064562419010290.
 19. Vokhmintcev, A., Melnikov, A., Pachganov, S., and Burlutskii, V. 2019. The New Combined Closed-Solution for 3D Reconstruction of Environment Based on Iterative Closest Point Algorithm. 7th Scientific Conference on Information Technologies for Intelligent Decision Making Support Proceedings. Ufa, Russia: 23–27.
 20. Langis, C., Greenspan, M., and Godin, G. 2001. The parallel Iterative closest point algorithm. *IEEE 3-rd International Conference 3D Digital Imaging and Modeling Proceedings*. Quebec City, Canada: 195–204.
 21. Thrun, S., Burgard, W., and Fox, D. 2000. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. *IEEE International Conference on Robotics and Automation Proceedings*. San Francisco, USA: 321–328.
 22. Vokhmintcev, A., Sochenkov, I., Kuznetsov, V., and Tikhonkikh, D. 2016. Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram. *Journal of Doklady Mathematics*. 93(1): 37–41. doi: 10.1134/S1064562416010178.
 23. *New York University depth dataset V2*. Available at: https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html (accessed December 28, 2012).

A.V. Vokhmintcev. Chelyabinsk State University, 129 Bratiev Kashirinykh st., 454001 Chelyabinsk, Russia, e-mail vav@csu.ru