

УДК 004.93

**МОДЕЛЬ ФОРМИРОВАНИЯ ИЗОБРАЖЕНИЙ ДЛЯ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ СЦЕН ПО ДАННЫМ ВИДЕОЭНДОСКОПИЧЕСКИХ ИССЛЕДОВАНИЙ**

*Е.А. ГОЛОВАТАЯ, канд. техн. наук, доц. В.С. САДОВ  
(Белорусский государственный университет, Минск)*

*В работе рассматриваются геометрические модели формирования изображений для последующей трехмерной реконструкции, а также предлагается модель широкоугольной сферической проекции для обработки изображений, полученных с использованием широкоугольных линз, например, из видеоэндоскопических систем.*

**Ключевые слова:** *трехмерная реконструкция, структура из движения, обработка видеоэндоскопических изображений.*

Большинство существующих методов трехмерной реконструкции сцен по изображениям этой сцены основываются на предположении о том, что каждое из изображений является некоторой двумерной проекцией точек трехмерного пространства на плоскость изображения при съемке. Это, в свою очередь, позволяет выполнить соответствующие геометрические построения и установить взаимосвязи между координатами точек на различных изображениях.

Таким образом, для корректной работы алгоритмов трехмерной реконструкции необходимо задать модель формирования проекции и модель ограничений при формировании проекции, которая позволяет связывать точки различных проекций между собой.

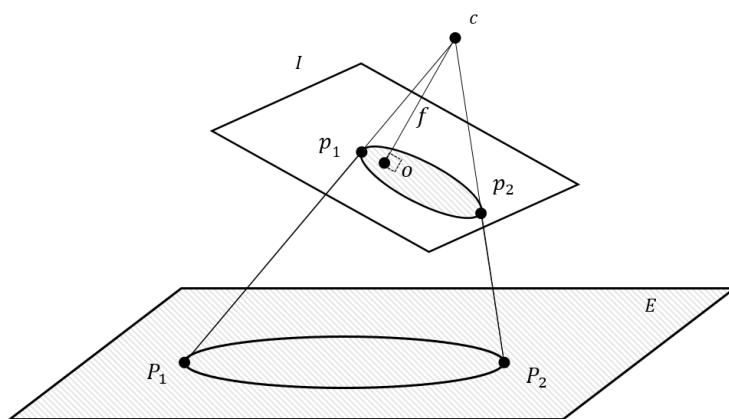
Традиционно в задачах фотограмметрии для моделирования самой съемки используется модель камеры-обскуры. Физически камера-обскура представляет собой фотографическую систему, в которой роль объектива исполняет небольшое отверстие. Световые лучи, проходя через отверстие, создают перевернутую проекцию исходной сцены на экранной плоскости, перпендикулярной нормали к отверстию. Изображение может быть зафиксировано с использованием фоточувствительного экрана непосредственным засвечиванием или с использованием цифровой матрицы на базе прибора с зарядовой связью (ПЗС) или комплементарной структуры металл-оксид-полупроводник (КМОП). Диаметр щели подбирается, как правило экспериментально, для того, чтобы обеспечить достаточную резкость изображения, формируемого на экранной плоскости: при слишком больших значениях радиуса изображения будет становиться заметным из-за расфокусировки поступающих лучей, а при слишком малых на изображении начнут проявляться дифракционные эффекты.

Конфигурация системы формирования проекции в камере-обскуре означает, что полученное изображение обладает двумя важными свойствами. Во-первых, в полученных изображениях полностью отсутствуют дисторсионные искажения: поскольку в системе не используются линзы, их кривизна не искажает изображение, и сформированная проекция ортоскопически в точности повторяет исходную сцену. Во-вторых, из-за относительных размеров отверстия, необходимых для обеспечения резкости проекции, полученное изображение обладает очень большой глубиной резкости, т.е. на нем не наблюдается размытия из-за удаления или приближения объектов съемки.

Практически во всех прикладных задачах для съемки используются камеры с намного более сложной оптической системой, чем камера-обскура. Это обусловлено, в первую очередь, тем, что при использовании камеры-обскуры для корректного формирования изображения требуется крайне высокое время выдержки – от секунды до нескольких суток. Таким образом, для самой задачи трехмерной реконструкции используется лишь соответствующая модель формирования изображения, т.е. использование такой модели для более современных оптических систем подразумевает обязательную калибровку для коррекции дисторсии и обеспечение полной глубины резкости.

Геометрическая модель формирования изображения с использованием камеры-обскуры – модель центральной проекции. Согласно этой модели проекцию трехмерной сцены из любой точки можно описать на основании параметров ориентирования камеры. Параметры ориентирования камеры разделяют на внешние и внутренние. Под внешними параметрами ориентирования подразумевают положение и ориентацию камеры в трехмерном пространстве. В глобальной системе координат, связанной с трехмерной сценой, внешние параметры ориентирования задаются шестью степенями свободы: тремя координатами положения камеры  $(x_c, y_c, z_c)$ , а также тремя углами ориентации камеры относительно осей глобальной системы координат  $(\omega_c, \varphi_c, \beta_c)$ . Под внутренними параметрами ориентирования понимают дополнительные характеристики самой камеры, которые влияют на формирование изображения. Самый главный из параметров внутреннего ориентирования – фокусное расстояние  $f$ , которое определяется как расстояние от центра проекции

до плоскости изображения вдоль оси, перпендикулярной плоскости изображения. Как правило, при фотограмметрической съемке захват изображения осуществляется с постоянным фокусным расстоянием, поэтому этот параметр считается известной константой. Согласно модели центральной проекции при формировании изображения некоторой точки  $P_i$  трехмерной сцены с координатами  $(X_i, Y_i, Z_i)$  в глобальной системе координат, проекция  $p_i$  этой точки может быть найдена следующим образом: из центра проекции камеры  $c$ , координаты которого в глобальной системе координат определяются координатами положения камеры  $(X_0, Y_0, Z_0)$ , строится прямая в эту точку трехмерного пространства, после чего определяется точка пересечения  $p_i$  этой прямой с плоскостью изображения  $I$ . При этом плоскость изображения удалена от центра проекции на величину  $f$  вдоль единичного вектора, сориентированного в пространстве относительно осей глобальной системы координат под углами  $(\omega, \varphi, \beta)$ . Пример формирования изображения для такой модели представлен на рисунке 1.



$c$  – центр проекции камеры,  $f$  – фокусное расстояние камеры,  $I$  – плоскость изображения,  
 $o$  – центр изображения,  $E$  – некоторый объект трехмерной сцены,

$P_1, P_2$  – некоторые точки трехмерной сцены,  $p_1, p_2$  – проекции точек  $P_1, P_2$  на изображении

Рисунок 1. – Геометрическая интерпретация формирования изображения в модели центральной проекции

Следует отметить, что в настоящей камере-обскуре центр проекции совпадает с положением щели, а плоскость изображения расположена за ней, т.е. центр проекции, в отличие от представленного изображения, располагается между объектом трехмерной сцены и плоскостью проекции. Можно показать, что для такого случая все геометрические построения остаются аналогичными. С другой стороны, большинство современных линзовых оптических систем обеспечивают положение центра проекции за плоскостью изображения.

В модели центральной проекции координаты любой точки на проекции в плоскости изображения однозначно определяются глобальными координатами соответствующей исходной точки трехмерной сцены и параметрами ориентирования камеры на момент получения снимка. Для этого необходимо рассмотреть произвольную точку исходной сцены с известным радиус-вектором  $\vec{P} = [X_p \ Y_p \ Z_p]^T$  в глобальной системе координат, радиус-вектор положения центра проекции камеры  $\vec{C}_0 = [X_0 \ Y_0 \ Z_0]^T$  в глобальной системе координат, и неизвестный вектор  $\vec{p}$ , построенный из центра проекции камеры  $\vec{C}_0$  в точку  $\vec{P}$ . Можно показать, что эти векторы образуют треугольник, т.к.  $\vec{P} = \vec{C}_0 + \vec{p}$ . С другой стороны, неизвестный радиус-вектор  $\vec{p}$  будет коллинеарен вектору  $\vec{p}_c = [x_p \ y_p \ z_p]^T$ , направленному из центра проекции в соответствующую точку на изображении и заданному в системе координат, связанной с изображением. Направление вектора  $\vec{p}_c$  в глобальной системе координат может быть определено умножением на матрицу поворота  $R$  [1], которая определяется на основании углов ориентации камеры в пространстве как

$$\begin{bmatrix} \cos(\varphi)\cos(\beta) + \sin(\varphi)\sin(\omega)\sin(\beta) & -\cos(\varphi)\sin(\beta) + \sin(\varphi)\sin(\omega)\cos(\beta) & \sin(\varphi)\cos(\omega) \\ \cos(\omega)\sin(\beta) & \cos(\omega)\cos(\beta) & -\sin(\omega) \\ -\sin(\varphi)\cos(\beta) + \cos(\varphi)\sin(\omega)\sin(\beta) & \sin(\varphi)\sin(\omega) + \cos(\varphi)\sin(\omega)\cos(\beta) & \cos(\varphi)\cos(\omega) \end{bmatrix}. \quad (1)$$

Таким образом, т.к. векторы  $\vec{p}$  и  $R\vec{p}_c$  коллинеарны, можно записать следующие зависимости:

$$\vec{P} = \vec{C}_0 + \vec{p} = \vec{C}_0 + mR\vec{p}_c \Leftrightarrow \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + mR \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}, \quad (2)$$

где  $(X_p, Y_p, Z_p)$  – трехмерные координаты исходной точки сцены в глобальной системе координат;  
 $(X_0, Y_0, Z_0)$  – трехмерные координаты центра проекции камеры на момент получения снимка;  
 $m$  – неизвестный коэффициент пропорциональности, введенный в силу коллинеарности векторов точки до и после проекции относительно центра проекции;  
 $R$  – матрица поворота, задаваемая с помощью углов ориентации  $(\omega, \varphi, \beta)$  по формуле (1);  
 $(x_p, y_p, z_p)$  – координаты проекции исходной точки в координатной системе изображения, связанной с центром проекции. Другими словами,  $(x_p, y_p)$  – координаты точки на изображении относительно точки пересечения основной оптической оси с плоскостью изображения,  $z_p$  – расстояние от центра проекции до плоскости изображения. Можно заметить, что расстояние  $z_p$  не зависит от положения точки и всегда равно фокусному расстоянию  $f$ .

Хотя соотношения (1), (2) позволяют установить взаимосвязь между координатами точки в глобальном пространстве и координатами ее проекции на изображении, полностью описать всю геометрию сцены не позволяет неизвестный масштабный коэффициент  $m$ . Этот масштабный коэффициент определяется расстоянием от произвольной точки в плоскости изображения до самого объекта в исходном трехмерном пространстве. С другой стороны, выражение (1) с учетом того, что  $R^{-1} = R^T$ , можно преобразовать относительно координат точки на изображении:

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \frac{1}{m} R^T \begin{bmatrix} X_p - X_0 \\ Y_p - Y_0 \\ Z_p - Z_0 \end{bmatrix}, \quad (3)$$

Выражение (3) позволяет составить систему из трех уравнений, из которых масштабный коэффициент может быть выражен через известную координату  $z_p$ . В результате преобразований можно получить пару уравнений для координат  $(x_p, y_p)$ :

$$x_p = x_0 + z_p \frac{r_{11}(X_p - X_0) + r_{21}(Y_p - Y_0) + r_{31}(Z_p - Z_0)}{r_{13}(X_p - X_0) + r_{23}(Y_p - Y_0) + r_{33}(Z_p - Z_0)}, \quad (4)$$

$$y_p = y_0 + z_p \frac{r_{12}(X_p - X_0) + r_{22}(Y_p - Y_0) + r_{32}(Z_p - Z_0)}{r_{13}(X_p - X_0) + r_{23}(Y_p - Y_0) + r_{33}(Z_p - Z_0)}, \quad (5)$$

где  $(X_p, Y_p, Z_p)$  и  $(x_p, y_p, z_p)$  – координаты точки объекта в глобальной системе координат и в системе координат изображения;

$(X_0, Y_0, Z_0)$  – координаты начала системы координат изображения в глобальной системе координат;

$r_{ij}$  – элементы матрицы  $R$ , определяемые из (1).

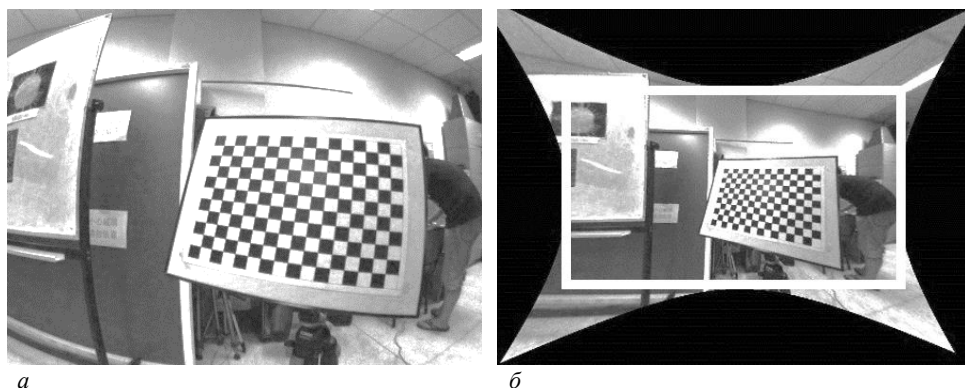
Уравнения (4), (5) называются уравнениями коллинеарности и описывают взаимосвязь между глобальной трехмерной системой координат и двумерной системой координат снимка с использованием шести параметров внешнего ориентирования (трех координат центра  $(X_0, Y_0, Z_0)$  и трех углов ориентации камеры  $(\omega, \varphi, \beta)$ , присутствующих в  $r_{ij}$ ) и трех параметров внутреннего ориентирования (фокусного расстояния  $f = z_p$  и двух координат точки  $(x_0, y_0)$  изображения, через которую проходит главная оптическая ось). Как правило, параметры внутреннего ориентирования определяются используемой оптической системой и могут считаться известными [1].

Уравнения коллинеарности (4), (5) могут быть построены для любых точек, для которых известны совпадающие координаты в глобальном пространстве, и объединены в общую систему уравнений. Например, если на одном изображении известны координаты трех точек в глобальной системе координат, для каждой из этих точек можно составить пару уравнений коллинеарности с шестью неизвестными параметрами внешнего ориентирования камеры, которая использовалась для получения снимка, и сформировать таким

образом систему из шести уравнений и с шестью неизвестными. Кроме того, при наличии нескольких изображений и координат одних и тех же точек пространства на этих изображениях, можно для каждого из них составить уравнения коллинеарности, которые и объединить в общую систему уравнений. Для достаточного количества точек количество уравнений превысит количество неизвестных, поэтому вместо прямого поиска решения этой системы можно ввести вектор ошибки проекции и осуществлять его минимизацию каким-либо численным методом оптимизации. После решения системы можно определить не только параметры внешнего ориентирования камеры, но и трехмерные координаты использованных точек, что может служить основой для последующей трехмерной реконструкции. Такой метод восстановления трехмерных координат точек по их изображениям на основании модели проекции называется методом связей.

Прямое решение системы, составленной из уравнений (4), (5), является достаточно сложной задачей, поэтому для поиска оптимального решения систему уравнений, как правило, линеаризуют и приводят к приближенной системе линейных алгебраических уравнений с использованием разложения в ряд Тейлора до первого дифференциального члена. Полученная матрица в линейном виде является в значительной степени разреженной. Это позволяет ввести в алгоритмы численной оптимизации дополнительные предположения и приближения для ускорения их работы. В частности, достаточно эффективной является реализация алгоритма оптимизации Левенберга – Марквардта с разбиением на блоки дополнением Шура, после чего две из четырех полученных подматриц могут быть дополнительно преобразованы с использованием разложения Холецкого [3]. Кроме того, эффективным является подход с использованием метода сопряженных градиентов и хранением значений матрицы в словарном виде.

Как уже упоминалось, одной из особенностей модели камеры-обскуры и основанной на ней модели центральной проекции является предположение о том, что в исходном изображении отсутствуют дисторсионные искажения. На практике это означает, что для использования соотношений (4), (5) и построенных на их основании систем уравнений для определения положений камеры в пространстве и последующей трехмерной реконструкции необходимо ввести способ учета дисторсионных искажений. Традиционный подход к этой проблеме подразумевает ручную или автоматическую калибровку камеры с использованием некоторой известной модели дисторсии таким образом, чтобы восстановить исходное изображение. Пример коррекции дисторсии с использованием ручной калибровки представлен на рисунке 2.



***а* – исходное изображение, *б* – изображение после коррекции [4]**  
**Рисунок 2. – Прямая коррекция дисторсии: белым выделена область, которая может использоваться для дальнейшей обработки**

Как видно из рисунка 2, существенным недостатком прямой коррекции дисторсии являются значительные нелинейные искажения исходного изображения при удалении от центра, особенно ближе к углам. Это обуславливает необходимость обрезки изображения после коррекции дисторсии, чтобы устранить негативное влияние перспективных искажений. Это, в свою очередь, приводит к значительной потере информации. Для сверхширокоугольных линз (с углом обзора до  $140^\circ$ ) количество пикселей, которое необходимо отбросить после коррекции дисторсии, может достигать 30% исходного изображения.

Одной из важных областей, в которых распространено использование таких видов линз, являются видеоскопические обследования [2]. Современное видеоскопическое оборудование, как правило, имеет достаточно сложную оптическую систему с широким углом обзора, чтобы обеспечить максимальную видимость для специалиста при проведении обследования. При проведении трехмерной реконструкции по таким данным потеря информации в результате коррекции дисторсии в значительной степени ухудшает качество реконструкции, что обуславливает необходимость разработки более сложной модели формирования проекции, в которой учитывается дисторсия.

В качестве альтернативы калибровке и коррекции дисторсии в данной работе предлагается модель сферической широкоугольной проекции для формирования изображения, которую можно использовать вместо модели центральной проекции. Кроме того, выпуклость линзы позволяет оптической системе захватывать световые лучи не только непосредственно до границ некоторой искусственной проекционной плоскости, но и в более широком диапазоне, что может быть удобнее для последующего сопоставления получаемых проекций в трехмерном пространстве. Графически эта модель представлена на рисунке 3.

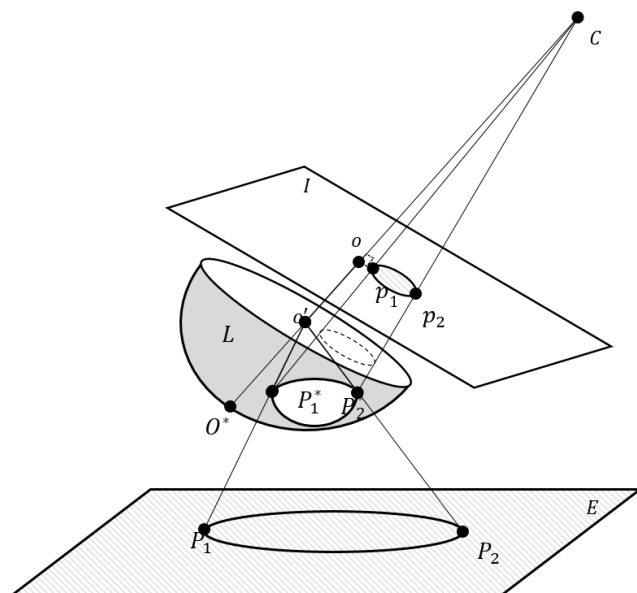


Рисунок 3. – Модель формирования изображения на основе сферической широкоугольной проекции

Как и в модели центральной проекции, камера  $c$  ориентирована вдоль главной оптической оси  $co$ , перпендикулярно которой расположена плоскость изображения. При этом для формирования самой проекции точек  $P_i$  сцены  $E$  они сначала проецируются на поверхность сферической линзы  $L$  через ее собственный центр проекции  $o'$  в точки  $P_i^*$ , после чего преломляются в соответствии с кривизной линзы на некоторый угол в зависимости от угла падения на линзу. Конечная проекция  $p_i$  осуществляется с поверхности линзы на плоскость  $I$  через центр проекции  $c$ . При незначительной кривизне поверхность линзы можно считать плоскостью; в этом случае модель становится эквивалентна модели центральной проекции.

В модели центральной проекции координаты проекции координаты точек в реальном пространстве связаны с координатами на изображении уравнениями коллинеарности (4), (5). Двумерная координатная система изображения  $(x_{ij}, y_{ij})$  определяет положение точки на изображении. При дополнительной проекции на сферическую поверхность эти координаты будут однозначно определять некоторую точку на поверхности сферы. Угол обзора линзы определяет расстояние  $h$  от края сферической поверхности до плоскости изображения, которая является сечением этой поверхности. Фокусное расстояние  $f$  позволяет при помощи модели центральной проекции связать точку на поверхности сферы с некоторой точкой в координатной системе плоского снимка, как показано на рисунке 4.

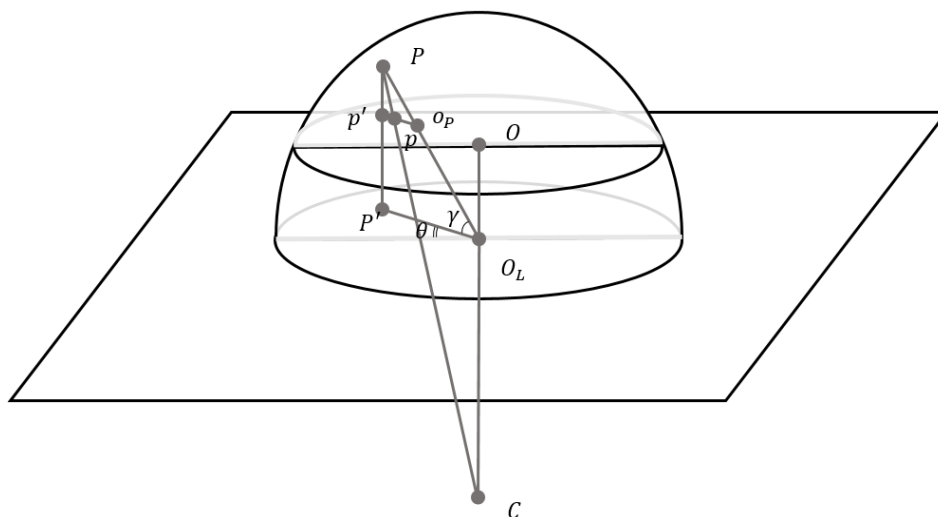
Если некоторая точка  $P^*$  в пространстве проецируется на поверхность сферы радиуса  $r$  в точке  $P$ , это означает, что в сферических координатах с центром в точке  $O_L$  центра сферы линзы местоположение точек  $P^*$  и  $P$  определяется одинаковыми значениями пары углов:  $\theta_p$  поворота в плоскости сечения, перпендикулярного главной оптической оси, и  $\gamma_p$  наклона в плоскости сечения, в которой лежит главная оптическая ось. Расстояние  $OO_L = h(\xi, r)$  от центра сферы линзы до центра проекционной плоскости известно и определяется углом обзора оптической системы  $\xi$  и радиусом линзы  $r$ . Расстояние  $O_L C = f$  от центра сферы линзы до центра проекции известно и определяется фокусным расстоянием. Проекция точки  $P$  с поверхности сферы на плоскость рисунка осуществляется через центральную проекцию с центром в точке  $C$ , т.е. точка проекции  $p$  лежит на пересечении прямой  $PC$  и плоскости изображения.

В двумерной координатной системе изображения положение точки определяется расстоянием  $\tau$  от центра изображения  $O$  до точки  $p$  и углом  $\theta$ . Геометрическими построениями можно получить выражения, связывающие двумерные координаты точки на проекции  $(x_p, y_p)$  со значениями углов  $(\gamma_p, \theta_p)$ :

$$x_{ij} = \frac{r(h+f)\cos\gamma_p\cos\theta_p}{r\sin\gamma_p - f}, \quad (6)$$

$$y_{ij} = \frac{r(h+f)\cos\gamma_p\sin\theta_p}{r\sin\gamma_p - f}. \quad (7)$$

С другой стороны, сами значения углов  $(\gamma_p, \theta_p)$  для этой точки в пространстве могут быть получены переводом ее трехмерных координат  $(X_p, Y_p, Z_p)$  в сферическую систему координат с началом координат в центре сферической поверхности линзы и основной осью, направленной вдоль главной оптической оси. Для соответствующих преобразований можно использовать формулу Эйлера – Родрига [5]. Выражения (6), (7) для произвольной точки в пространстве являются аналогами уравнений коллинеарности (4), (5) для модели сферической широкоугольной проекции. Таким образом, соответствующие уравнения могут использоваться для формирования модели ограничений при последующей оптимизации для задачи трехмерной реконструкции. В отличие от уравнений коллинеарности в модели центральной проекции, аналитическое определение частных производных выражений углов для линеаризации системы уравнений разложением в ряд Тейлора может быть затруднительно для приведенного вида функций. В качестве решения этой проблемы все описанные функции могут быть реализованы программно для поддержки вычисления в произвольной точке, после чего значения частных производных по определенному аргументу могут быть вычислены с использованием простой трехточечной разностной схемы, т.к. для всех точек и их окрестностей, участвующих в полученных выражениях, эти производные существуют. Кроме того, фиксированные параметры модели  $h, r, f$  можно также положить неизвестными при наличии достаточного количества снимков.



**Рисунок 4. – Построение проекции точки с поверхности сферической линзы в модели широкоугольной сферической проекции**

Модель сферической проекции позволяет учитывать координаты точек для каждого пикселя полученной проекции, а также их пространственное расположение. При формировании матрицы ограничений в линейной форме она имеет схожий разреженный вид с матрицей метода связей для модели центральной проекции, поэтому для поиска оптимального решения могут применяться оптимизации метода Левенберга-Марквардта, описанные ранее.

Таким образом, разработанную модель сферической широкоугольной проекции можно использовать для формирования первичных ограничений при пересчете трехмерного представления сцены на основании метода связей. Учет всех точек полученного изображения позволяет учитывать дисторсионные

искажения проекции на уровне самой модели и использовать эту информацию для более точного восстановления окружающих сцен. В качестве одной из важных областей применения рассмотренной модели можно выделить задачу трехмерной реконструкции на основании данных, полученных с использованием видеоэндоскопических систем. Представленная модель и уравнения для метода связок могут использоваться для оценки положений совпадающих точек в глобальном трехмерном пространстве моделирования и построения на их основе разреженного облака точек, которое в дальнейшем может уплотняться и интерполироваться в полигональные поверхности для формирования трехмерного изображения исследуемых внутренних органов на основании изображений и видеоряда.

#### ЛИТЕРАТУРА

1. Close-Range Photogrammetry and 3D Imaging / T. Luhmannx [et al.]. – 2nd ed. – Berlin : De Gruyter, 2013. – 684 p.
2. Dremel, H.W. General Principles of Endoscopic Imaging / H.W. Dremel // Principles and Practice of Interventional Pulmonology ; ed.: A. Ernst, F. Herth. – New York : Springer, 2013. – P. 15–25.
3. Lourakis, M. I. A. SBA: A Software Package for Generic Sparse Bundle Adjustment / M.I. A. Lourakis, A.A. Argyros // ACM Trans. Math. Software, ACM, New York, NY, USA. – 2009. – Vol. 36, № 1. – P. 1–30.
4. Radial distortion correction from a single image of a planar calibration pattern using convex optimization / Xianghua Ying [et al.] // IEEE International Conference on Image Processing (ICIP), IEEE. / – Paris, 2014. – P. 3440–3443.
5. Dai, J.S. Euler–Rodrigues formula variations, quaternion conjugation and intrinsic connections / J.S. Dai // Mechanism and Machine Theory, International Federation for the Theory of Machines and Mechanisms // Elsevier. – 2015. – Vol. 92. – P. 144–152.

*Поступила 09.09.2019*

#### MODEL OF IMAGE ACQUISITION FOR 3D SCENE RECONSTRUCTION FROM VIDEOENDOSCOPIC IMAGING DATA

*K. HALAVATAYA, V. SADAU*

*The paper discusses geometric image formation models for 3D reconstruction, and proposes a wide-angle spherical projection model to work with images acquired using wide-angle lenses, like those used in video endoscopic imaging.*

**Keywords:** *3D reconstruction, structure from motion, videoendoscopic image processing.*