

Комбинированный метод вычисления значения диспаратности на стереоизображениях в задачах стереодальнометрии

А.Н. Волкович

Объединенный институт проблем информатики НАН Беларуси, 220012, ул.Сурганова 6, Минск, Республика Беларусь

Аннотация

В статье рассматривается решение задачи восстановления трехмерной информации на основе стереоизображений. Предлагается оригинальный комбинированный подход к вычислению диспаратности, а также приводится вариант решения проблемы разнородности исходных данных при вычислении фактических метрических параметров.

Ключевые слова: стереоизображения; карты диспаратности; операторы градиентов; спутниковые фотоснимки; дальнометрические системы; оптические системы

1. Введение

Задачи сопоставления и поиска изображений являются одними из основных задач в компьютерном зрении. Качество поиска, низкая чувствительность к искажениям – основные требования, предъявляемые к алгоритмам. Существует множество подходов для их решения, а также широкий спектр технических решений задач дальнометрии, использующих лазерные и иные системы. Вместе с тем ряд задач предполагает невозможность применения активных дальнометрических систем. В описываемом проекте планируется разработка и реализация эффективных методов поиска участков изображений на основе гистограммных дескрипторов характеристик локальных окрестностей пикселей.

Актуальность данного проекта обеспечивается в первую очередь необходимостью развития методов решения трудоемких вероятностно-геометрических задач цифровой обработки случайных полей и изображений, объем и сложность которых, непрерывно возрастают в связи с развитием информационно-вычислительных технологий и потребностями научно-технического прогресса. Кроме того, специализированная направленность новых методов исследования на эффективное решение конкретных прикладных задач сочетается при этом с модульным принципом построения и универсальностью программных компонентов всех создаваемых систем, что позволит успешно применять их при исследовании проблемных и прикладных вопросов не только при обработке цифровых изображений, но и в смежных дисциплинах.

2. Современное состояние и особенности применения дальнометрических систем

Задача определения расстояния до объекта является крайне актуальной в таких областях как геодезия, военное дело, навигация и компьютерное зрение. Для определения расстояния на практике используются дальномеры — устройства, предназначенные для определения расстояния от наблюдателя до объекта.

Дальномерные приспособления делятся на активные и пассивные. Первоначальное развитие дальнометрических приспособлений шло среди пассивных систем, однако в последнее время наибольшее распространение получили активные дальномеры ввиду простоты их реализации, неприхотливости использования и достаточно высокой точности замеров.

Принцип действия дальномеров активного типа состоит в измерении времени, которое затрачивает посланный дальномером сигнал для прохождения расстояния до объекта и обратно. Скорость распространения сигнала (скорость света или звука) считается известной. Также существуют активные дальномеры, производящие вычисление расстояния, основываясь на оценке изменений параметров отраженного сигнала, например, фазы или мощности.

Следует отметить, что существует ряд задач, в которых применение активных дальномеров затруднено либо не возможно. К таким задачам можно отнести применение дальномеров на малозаметных платформах, использование излучателей на которых, приводит к их демаскировке, а также на больших дистанциях, требующих установки излучателей с мощностями, несущими опасность.

Измерение расстояний пассивными дальномерами основано на определении высоты h равнобедренного треугольника ABC , например по известной стороне $AB = l$ (базе) и противолежащему острому углу b (т. н. параллактическому углу). При малых углах β (выраженных в радианах) $h = l/b$. Одна из величин, l или b , обычно является постоянной, а другая — переменной (измеряемой). По этому признаку различают дальномеры с постоянным углом и дальномеры с постоянной базой. В общем случае, основываясь на расстоянии между точками наблюдения l (база) и угол смещения α , рассчитывается расстояние до объекта:

$$h = \frac{l}{2 \sin \frac{\beta}{2}}$$

К сожалению, до недавнего времени использование стереовизионных принципов в дальнометрических системах практически не рассматривалось ввиду относительной сложности реализации и низкого качества цифровых

изображений. Однако в последние годы качественные параметры съемочного оборудования значительно возросли, что позволяет говорить о возможности разработки систем такого рода. Кроме того, стереовизионная фотодальнометрическая система может обладать помимо высокой малозаметности (ввиду реализации пассивной методики вычисления) также функционалом, позволяющим производить постсъёмочное получение дальнометрической информации ввиду возможности сохранения стереопар и повторного вычисления параллакса для множества точек изображения в зоне перекрытия и резко отображаемого пространства снимка – в отличие от систем, совмещенных с активными дальномерами, позволяющими получить дальнометрическую информацию только в момент съемки-замера.

Основываясь на изложенных выше фактах, видится крайне актуальной разработка как теоретических методов в области стереодальнометрии, так и технических программно-аппаратных решений.

3. Пассивные системы дальнометрии

Цифровое изображение, полученное при помощи пассивной стереосистемы, несет в себе лишь цветовую или яркостную информацию и не обладает какими-либо вспомогательными данными.

Поскольку изображения реального мира включают в себя достаточно узкий и однотипный набор цветов или яркостей, при решении задачи определения сопряженных точек речь идет об отождествлении не отдельных точек, а фрагментов изображений, лежащих в окрестностях этих точек. Таким образом, крайне важно знать, если на одном изображении взята какая-то точка, то где на втором изображении находится ей сопряженная. И каким образом сопоставить эти фрагменты.

Сопоставление окрестностей сопряженных точек не поддается строгой формализации, поскольку в его основе лежит проблема идентификации по изображениям фрагментов трехмерного мира, которые с трудом поддаются адекватному формальному описанию. Возникающие при съемке и зависящие от ракурса проективные и яркостные искажения приводят к тому, что в изображениях одних и тех же участков поверхности, снятых под разными ракурсами, могут появиться значительные отличия. Принципиально важно то, что эти отличия зависят не только от геометрии съемки, но и от геометрических и физических характеристик самой поверхности. Расположение источника света по отношению к поверхности влияет на распределение освещенности. Положение элементов поверхности, их отражательные и рассеивающие свойства определяют количество энергии, попадающей в объективы камер, а, следовательно, и локальные отличия в яркости сопряженных фрагментов изображений.

Величина различий зависит от разницы в ракурсах съемки. Как правило, чем больше эта разница (в частности, чем больше база), тем менее похожими становятся изображения. Поэтому все методы сопоставления окрестностей сопряженных точек в большей или меньшей степени опираются не на формальный подход, а на изучение характера изображений, на возможность их предварительной обработки, на приведение к эпиполярной стереопаре, на построение эффективных по точности и быстродействию описателей окрестностей сопряженных точек для сопоставления.

В идеализированной ситуации значения функции схожести в процессе сканирования вдоль строки должен представлять точечный выброс для искомого пикселя при возврате нулевого значения подобия для всех остальных пикселей (окрестностей) строки.

Безусловно, при работе с реальными графическими данными такая комбинация возвращаемых функцией сходства значений невозможна. В свою очередь, при обработке исходных данных, обладающих достаточным количеством информации для идентификации локального региона, на графике функции сохраняется достаточно явный экстремум, позволяющий идентифицировать искомым пиксел изображения.

Таким образом, главной задачей при построении карты диспаратности является подбор варианта сравнения регионов, при котором экстремум значений функции подобия будет наиболее выраженным, что, соответственно, предполагает определение для точки некоторых характеристик, которые бы уникальным образом характеризовали точку изображения – так, чтобы парная точка на опорном изображении обладала идентичными или максимально близкими значениями аналогичных характеристик.

Финальным этапом вычисления диспаратности является агрегирование суммированных или усредненных стоимостей. При агрегировании в качестве итоговой диспаратности для точки p базового изображения выбирается такое d , при котором достигается минимальное значение стоимости. В реальной ситуации возможно нахождение нескольких значений d с одинаковыми и минимальными значениями стоимости (особенно при проведении усреднения стоимостей). Проблема возможной неоднозначности возникает у большинства методов, построения карт диспаратности, что во многом связано с оптическими, механическими, электронными особенностями камер. Решением многозначности может быть введение некоторых условий на значение диспаратности, например, наибольшее, среднее или наименьшее из возможных.

4. Комбинированный метод стереовосстановления

Наиболее часто на практике используются меры на основе суммы абсолютных разностей или суммы квадратов разностей. Обе функции (суммирование производится по заданному окну) позволяют производить вычисление стоимостей достаточно эффективно в случае, когда задается неявное предположение о том, что для данной точки базового изображения, соответствующий пиксел парного изображения имеет максимально близкое значение интенсивности. Таким образом, данные функции являются чрезвычайно чувствительными к качеству и параметрам

исходных данных (экспозиция, бликование, переэкспонирование, недоекспонирование, шум матрицы, случайные выбросы).

В процессе проведения вычислительного эксперимента на изображениях реального мира был определен факт того, что корреляционные методы сопоставления локальных участков изображений дают стабильные результаты на текстурированных участках и крайне низкую точность в однородных областях (не имеющих явных контрастных переходов). Обратившись к результатам исследований живых систем, можно сделать вывод о том, что в природе также плохо локализуется дистанция до однородного объекта и производится его привязка к границам за счет рефлекторных саккадических движений.

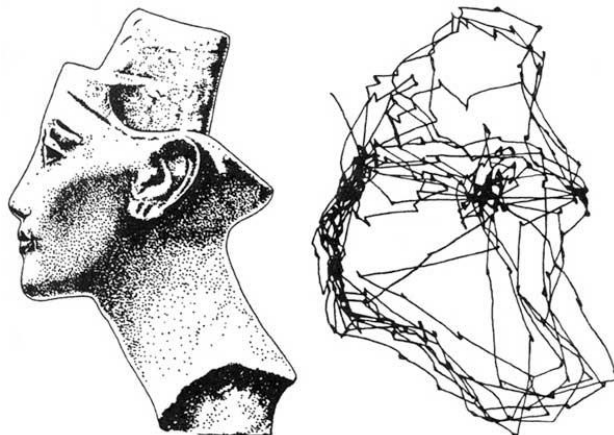


Рис. 1. Запись движения глаза (сканирование при рассматривании объекта изображения - головы Нефертити) по Ярбусу, 1965 г.

Таким образом, можно утверждать, что требуется комплексный подход к задаче стереодальнометрии, включающий в себя как рассмотрение возможностей увеличения уникальности средств идентификации, так и выработку комбинированного подхода с применением различных методик к окрестностям точек в зависимости от их локальных характеристик.

Работы с цифровыми стереоизображениями позволили определить следующую комбинированную методику: к участкам изображений, обладающим в своей площади контрастными объектами (перепады яркости), следует применять корреляционные методы с максимальной размерностью уникализации точек, а к однородным – производить привязку к удаленным контрастным переходам за счет построения совокупности векторов удаления по нескольким направлениям.

Таким образом, на этапе предобработки становится важным составление карты вычислений на основании близости к контрастным участкам. В целях классифицирования точки как находящейся на перепаде яркости изменение яркости, ассоциированное с данной точкой, должно быть существенно большим, чем изменение яркости в точке фона. В связи со спецификой локальных вычислений способ определения «существенных» значений состоит в установлении порога. В свою очередь для количественного выражения изменения яркости используют понятия первой и второй производных.

Определение точки изображения как точки перепада происходит в том случае, если ее двумерная производная первого порядка превышает некоторый заданный порог. Вычисление первой производной цифрового изображения основано на различных дискретных приближениях двумерного градиента. Направление вектора градиента совпадает с направлением максимальной скорости изменения функции f в точке (x, y) .

$$\begin{matrix} z1 & z2 & z3 \\ z4 & z5 & z6 \\ z7 & z8 & z9 \end{matrix}$$

Вычисление градиента изображения состоит в получении величин частных производных $G_x = df/dx$ и $G_y = df/dy$ для каждой точки. Один из способов нахождения первых частных производных G_x и G_y в конкретной точке состоит в применении следующего градиентного оператора Собеля:

$$\begin{aligned} G_x &= (z7 + 2 * z8 + z9) - (z1 + 2 * z2 + z3) \\ G_y &= (z3 + 2 * z6 + z9) - (z1 + 2 * z4 + z7) \end{aligned}$$

Для оператора Собеля, который выявляет горизонтальные и вертикальные контуры (перепады яркости), следует определить соответствующие маски для свертки с исходным изображением. Также можно изменить приведенные формулы таким образом, чтобы они давали максимальный отклик для контуров, направленных диагонально. Дополнительные пары масок оператора Собеля, предназначенные для обнаружения разрывов в диагональных направлениях можно определить как:

$$\begin{matrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \\ -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{matrix}$$

для точек, лежащих на диагональном перепаде -45 градусов;

для точек, лежащих на диагональном перепаде +45 градусов.

У каждой из масок сумма коэффициентов равна нулю, таким образом, данные операторы будут давать нулевой отклик на областях постоянной яркости, что характерно дифференциальному оператору. Рассмотренные маски применяются для получения составляющих градиента G_x и G_y . Для вычисления величины градиента эти составляющие необходимо использовать совместно, определяя величину градиента приближенно через абсолютные значения частных производных:

$$Gradient = |G_x| + |G_y|$$

На основании полученной карты градиентов строится карта вычислений. Данный процесс подразумевает признание области как малотекстурированной в том случае, если в окне поиска вокруг точки менее 15% площади занимают контрастные переходы.

Непосредственно вычисление диспаратности происходит в несколько этапов. На первом этапе обрабатываются высокотекстурированные участки с помощью корреляционных функций мер сходства, таких как, например, евклидово расстояние, кросс-корреляция и др.

В мировой практике при работе с изображениями в задачах стереовосстановления обычно используется только информация о яркости как критерии сравнения точек изображений. Недостатком данного подхода является множественность интерпретаций цветов для точек с одинаковым значением яркости. Кроме того, следует принять во внимание факт неравномерности восприятия цветного и монохромного изображения. Данная особенность учитывается в методах деградации цветовой модели изображения к 256 оттенкам серого за счет введения коэффициентов, применяемых к соответствующим каналам.

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

Исходя из того, что большинство изображений изначально формировалось цветным датчиком в цветном виде, в целях повышения эффективности очевидным видится переход к работе с цветовой информацией.

При переходе к работе с тремя компонентами изображения можно представить в виде "облака" точек в трехмерном пространстве с осями, соответствующими цветовым каналам изображения. Однако, RGB пространство не является ортогональным, что обусловлено спецификой зрительного анализатора человека, имеющего различное число палочек и колбочек, восприимчивых к тому или иному цвету.

Поскольку мерой сравнения точек в трехмерном пространстве выступают корреляционные функции, тем или иным путем использующие евклидово расстояние, которое корректно рассчитывается в ортогональных системах, следует производить ортогонализации пространства RGB в пространство XYZ.

Представление базисных цветов RGB, согласно рекомендациям ИТУ, в пространстве XYZ имеет следующие коэффициенты поправки:

$$\text{Красный: } x = 0,64 \quad y = 0,33$$

$$\text{Зеленый: } x = 0,29 \quad y = 0,60$$

$$\text{Синий: } x = 0,15 \quad y = 0,06$$

Таким образом, систему преобразований для перевода цветов между системами RGB и XYZ можно представить в следующем виде:

$$X = 0,431 * R + 0,342 * G + 0,178 * B$$

$$Y = 0,222 * R + 0,707 * G + 0,071 * B$$

$$Z = 0,020 * R + 0,130 * G + 0,939 * B$$

После приведения пространств к точкам могут применяться операции, справедливые для ортогональных систем. В свою очередь переход к «цветной» обработке изображений в 1,72 раза увеличивает потенциальную уникальность точки (максимальное расстояние между значениями яркости в градациях серого составляет 255 единиц, а между значениями цветов 416 единиц).

После выполнения данного шага производится определение расстояний от точек в малотекстурированных областях до ближайших контуров по нескольким направлениям. Таким образом формируется группа многомерных векторов характеристики, к которым применяется аналогичный подход, как и к векторам с цветовой информацией.

По мере выполнения этапов происходит заполнение карты диспаратности. По причине того, что все операции идут в строгом соответствии с картой вычислений, происходит автоагрегация результатов различных этапов в единую карту.

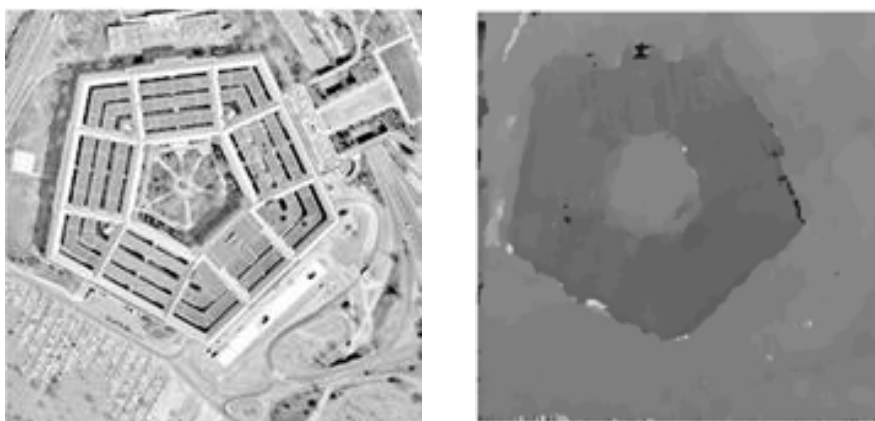


Рис. 2. Изображение стереопары и карта диспаратности.

Несмотря на многоэтапность выполнения, данный алгоритм обладает большим числом циклических стереотипных локально независимых операций, что дает возможность говорить о возможном распараллеливании алгоритма.

5. Расчет расстояния до объекта на основе разноразмерных исходных данных

Обобщенно принцип определения положения точек в пространстве на основании данных о диспаратности многократно описан в литературе. Положим, две камеры L и R установлены таким образом, что их оси X коллинеарны, а оси Y и Z , в свою очередь, параллельны. Центры камер смещены относительно друг друга на величину b , соответствующую базе стереоскопической системы. При наблюдении некоторой точки пространства P на левом изображении формируется точка P_l , а на правом – P_r .

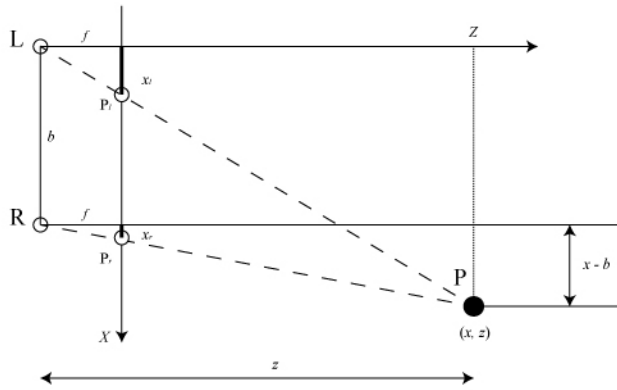


Рис. 3. Геометрическая модель стереоскопической системы.

Учитывая подобие двух пар треугольников, получим уравнения:

$$\frac{z}{f} = \frac{x}{x_l} \quad \frac{z}{f} = \frac{x - b}{x_r} \quad \frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}$$

Следует отметить, что по построению координаты точек изображений y_l и y_r могут считаться одинаковыми, что соответствует ректифицированной системе с жесткой связью между фотосистемами (3D-камера, зрительная система человека). Учитывая данное свойство, возможно преобразование системы уравнений для явного выражения координат x, y, z точки P в реальном пространстве на основе данных о координатах проекций точек на изображениях стереопары:

$$z = fb / (x_l - x_r) \quad x = \frac{x_l z}{f} = b + \frac{x_r z}{f} \quad y = \frac{y_l z}{f} = \frac{y_r z}{f}$$

Решение системы уравнений позволяет однозначно вычислить положение точки в пространстве.

К сожалению, приведенная выше система уравнений малоприменима при цифровом восстановлении, ввиду того, что происходит смешение различных систем размерности: значение диспаратности в пиксельном расстоянии, фокусное расстояние и база в метрических единицах. Однако возможно рассчитать дистанцию до точки, используя значение диспаратности, угол раскрытия створа объектива и базу стереосистемы.

Оптическую систему можно представить геометрически в виде треугольника: A – точка наблюдения; BC – изображение; α – горизонтальный угол раскрытия объектива.

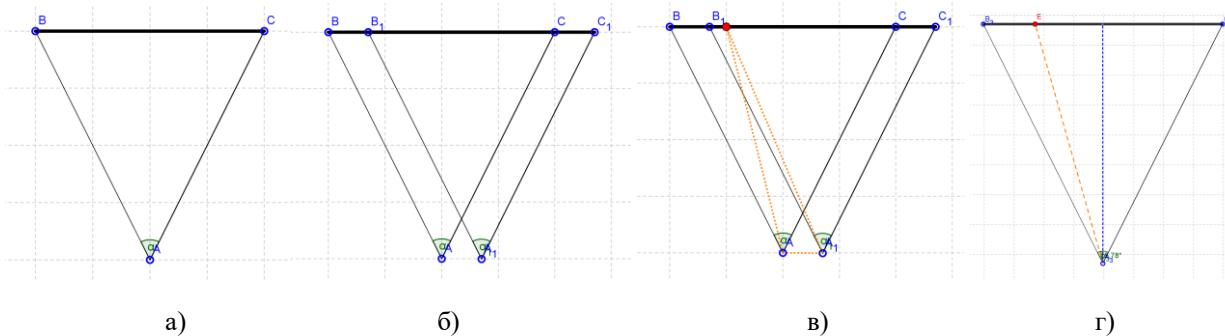


Рис.4. Геометрическая модель расчета дистанции по раскрытию створа объектива.

Таким образом, систему стереовидения можно представить в следующем виде: A, A_1 – точка наблюдения; BC – левое изображение; B_1C_1 – правое изображение; BC_1 – зона перекрытия; α – горизонтальный угол раскрытия объектива, а вычисление расстояний до объекта представляет собой задачу решения треугольника B_1AA_1 .

В рамках системы известно основание треугольника – база стереосистемы. Углы при основании могут быть вычислены через угол раскрытия объектива.

Расчет угла визирования на объект интереса для левой камеры выполняется следующим образом (Рис. 4в):

$$\angle EAA_1 = \beta = 90 \pm \arctg\left(\frac{\operatorname{tg} \frac{\alpha}{2} * \left(\frac{BC}{2} - BE\right)}{\frac{BC}{2}}\right)$$

Где: β – угол искомого треугольника; α – угол горизонтального раскрытия объектива; BE – координата X точки изображения; BC – ширина изображения (Рис 4г). Знак «+» используется в системе при $BE < \frac{BC}{2}$, «-» при $BE > \frac{BC}{2}$ соответственно, а при $BE = \frac{BC}{2}$ принимаем $\angle EAA_1 = \beta = 90$

В свою очередь, расчет угла визирования производится аналогично с поправкой на то, что знак «-» используется в системе при $BE < \frac{BC}{2}$, «+» при $BE > \frac{BC}{2}$ соответственно и при $BE = \frac{BC}{2}$ принимаем $\angle EA_1A = \beta_1 = 90$.

Получение третьего угла возможно по формуле:

$$\beta_2 = 180 - \beta - \beta_1$$

Далее по теореме синусов возможно определение длин сторон A_1E и AE

$$AE = AA_1 \frac{\sin \beta}{\sin \beta_2} \text{ и } A_1E = AA_1 \frac{\sin \beta_1}{\sin \beta_2}$$

В связи с возможным наклоном AE относительно горизонтальной плоскости системы необходимо приведение EA в плоскость системы

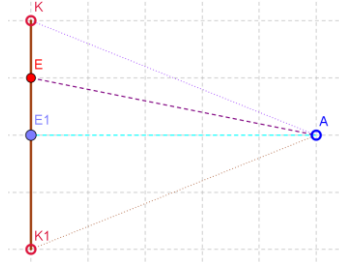


Рис. 5. Геометрическая модель вертикального склонения системы.

Если $KE < KK_1/2$

$$\vartheta = \arctg\left(\frac{\operatorname{tg} \frac{\alpha}{2} * \left(\frac{KK_1}{2} - AE\right)}{\frac{KK_1}{2}}\right)$$

Если $KE > KK_1/2$

$$\vartheta = \arctg\left(\frac{\operatorname{tg} \frac{\alpha}{2} * \left(AE - \frac{KK_1}{2}\right)}{\frac{KK_1}{2}}\right)$$

Если $BE = BC/2$

$$\begin{aligned} \vartheta &= 0 \\ \vartheta &= \angle EAE_1 \\ AE_1 &= AE \cos \vartheta \end{aligned}$$

Таким образом получено расстояние от опорной (левой) камеры до объекта и угол относительно основания (плоскости матриц) системы.

Следует отметить, что при увеличении значения отношения измеряемого расстояния к стереобазе повышается чувствительность системы к точности юстировки системы и качеству изображений, т.к. углы при основании системы принимают большие значения близкие к 90° , при которых крайне динамично меняются значения тригонометрических функций.

6. Заключение

В ходе проведения исследования автором было произведено изучение существующих методов обработки стереоизображений в задачах стереовосстановления, выявлены закономерности функционирования алгоритмов и определены причины их нестабильной работы. Предложена комбинированная методика обработки изображений, учитывающая характеристики локальных участков изображений. Кроме того, автором рассмотрена проблема разнородности исходных данных, необходимых для получения метрической информации о трехмерных объектах в области интересов, а также предложено ее решение.

Описанный выше алгоритм был реализован автором в виде программной библиотеки, которая в дальнейшем может быть использована в широком спектре приложений, т.к. матрица расстояний до точек изображения или до одной точки может быть переведена в конкретную систему координат, необходимую той или иной программной или программно-аппаратной системе. Также следует отметить, что организация возможности доступа пользователя к функциям, локализирующим различные этапы выполнения алгоритма, позволяет более гибко использовать библиотеку в приложениях, не требующих выполнения всего процесса стереообработки для решения поставленных задач.

Кроме того, описанная методика нашла свое применение в ряде программных и программно-аппаратных разработок, выполняющихся в Объединенном институте проблем информатики НАН Беларуси. В частности, в виде элемента мобильной топогеодезической системы и в виде программной библиотеки элементов системы ДЗЗ, разрабатываемой в рамках программы «Мониторинг СГ» Союзного государства.

Литература

- [1] Borodach, A. Automatic determination of matching points on two images / A. Borodach, A. Tuzikov // Proceedings of the 9th International Conference “Pattern Recognition and Information Processing”. 22-24 May, 2007, Minsk, Belarus. Vol. 1. P. 49-53.
- [2] Шапиро, Л. Компьютерное зрение / Л. Шапиро. — М.: БИНОМ. Лаборатория знаний, 2006. — 752 с.
- [3] Волкович, А.Н. Использование цветовых характеристик при построении карт диспаратности / А.Н. Волкович // Материалы международного конгресса РОПИ-2011. Нижний Новгород: ННГУ, 2011. С. 112-117. 64
- [4] Ляховский, В.В. Система автоматической реконструкции трехмерных сцен по нескольким изображениям / В.В. Ляховский, А.Н. Волкович, Д.В. Жук, А.В. Тузиков // Материалы V Белорусского космического конгресса, 25–27 октября 2011 г. – Минск, 2011. – Т. 2. – С. 129–133.
- [5] Жук, Д.В. Реконструкция трехмерной модели по двум цифровым изображениям / Д.В. Жук, А.В. Тузиков // Информатика. – 2006. – № 1. – С. 16-26.
- [6] Шульговский, В.В. Основы нейрофизиологии. [Электронный ресурс]. — Режим доступа: <http://www.braintools.ru/rubric/information/from-books/fundamentals-of-neurophysiology> (01.02.2017).