

Садыхов Р.Х., Ламовский Д.В.

ДЕТЕКЦИЯ СТЕРЕОЛИЦ В СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ

Введение. На сегодняшний день 3D моделирование лица человека находит очень широкое применение в системах обработки данных. Основные сферы применения такого рода объектов: анимирование персонажей в компьютерных играх и мультимедиа приложениях, высококачественная 3D анимация, идентификация человека в системах контроля доступа и видеонаблюдения и др. Для построения таких моделей в большинстве случаев используются специализированные установки, использующие набор (иногда до нескольких десятков) камер, источники структурированной подсветки или специализированные лазерные сканеры. В более простых случаях, когда точность модели не критична, она может строиться на базе статических изображений или кадров видеопоследовательности [2]. Иногда для построения таких грубых моделей требуется дополнительное вмешательство оператора [3]. Наряду с высоким качеством получаемых моделей, указанные подходы имеют недостатки. Они не применимы в случаях, когда модель лица должна быть получена в реальном масштабе времени, когда нет возможности управлять условиями получения изображений, а также когда невозможно вмешательство оператора.

Для систем, использующих 3D модели лица, для идентификации человека необходимо обеспечить высокое быстродействие и минимальные требования к положению объекта (лица) по отношению к устройству захвата. На рынке информационных технологий существуют подобные решения. Так, компания A4vision среди прочего биометрического оборудования создает устройства считывания трехмерной поверхности лица. Устройство представляет собой специализированный сенсор, интегрированный в один корпус с источником невидимого инфракрасного излучателя и ЖКИ дисплеем. Излучатель необходим для формирования невидимой структурированной подсветки. Дисплей используется для привлечения внимания сканируемого человека и выдачи ему указаний, с целью правильного позиционирования его головы перед сенсором. Производитель сообщает, что время идентификации по полученной модели составит не более 1 секунды. Стоимость устройства порядка 5000\$. Несмотря на заявленные хорошие показатели качества работы устройства, указанная технология все же требует от идентифицируемого человека выполнения определенных действий, что не всегда приемлемо.

В связи с широким распространением систем видеонаблюдения, использующих видеокамеры, все более высокого разрешения, возникает вопрос об использовании такого вида оборудования для получения информации о трехмерной поверхности лица. Решение такой задачи может быть выполнено с помощью двух подходов: с использованием одноканальных и многоканальных изображений. В первом случае для получения приемлемой модели необходимо точное позиционирование лица в фас. При этом полученная модель не может претендовать на высокую точность воспроизведения оригинальных параметров лица. Для более точного восстановления поверхности лица можно использовать изображения, полученные с нескольких камер, установленных под разными ракурсами. При этом для получения дополнительной информации могут быть применены методы стереовидения [6, 7].

Первым этапом при обработке изображений является выделение объектов интереса. В случае идентификации по портрету человека речь идет о детекции лиц на имеющихся изображениях. В данной работе мы представляем алгоритм детекции и сопоставления лиц на кадрах последовательности стереоизображений. Основной нашей целью являлось получить изображения лиц людей, проходящих в поле зрения нескольких камер систем видеонаблюдения. Лицо

человека в данном случае представляется в виде набора изображений, соответствующих различным ракурсам зрения. Основными требованиями к данной системе были: число лиц от 1 до 5, высокая скорость обработки кадров, размер лица не менее 100 на 100 пикселей для кадров размера 1024*768.

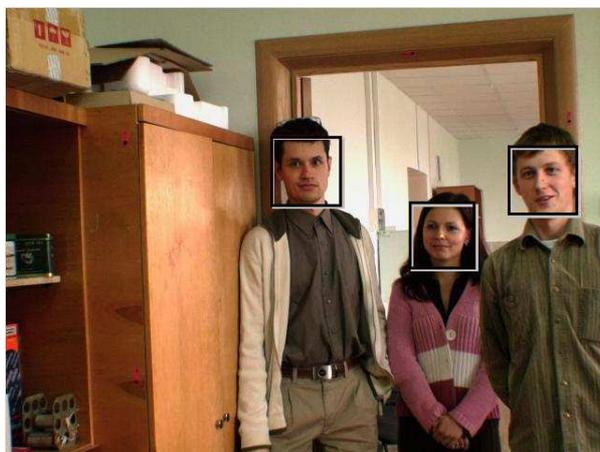
Калибровка камер. Классически калибровка камер в стереосистеме требует вычисления относительного положения оптических центров камер. В нашем случае такая информация не представляет интереса. Для решения имеющейся задачи необходимо определить фундаментальную матрицу, которая позволила бы вычислять соответствующую эпиполярную линию для любой точки кадра. То есть линию, на которой на втором кадре пары расположена точка, соответствующая выбранной. Эпиполярная геометрия в значительной степени уменьшает пространство поиска соответствующих точек между двумя стереокадрами. Для вычисления фундаментальной матрицы был использован восьмиточечный алгоритм [4] совместно с методом исключения выбросов RANSAC [5]. При этом для вычисления калибровки необходимо задать как минимум восемь пар соответствующих друг другу точек на стереокадрах. Полученную фундаментальную матрицу можно использовать для вычисления эпиполярных линий с помощью следующего выражения:

$$l = Fx, \quad (1)$$

где x – точка на одном из кадров стереопары, F – фундаментальная матрица, а l – соответствующая эпиполярная линия.

Детекция стереолиц. Для получения стереоизображений лиц, которые представляют собой наборы изображений лиц, полученных с различных ракурсов, мы используем разреженный алгоритм стереосопоставления. Указанный алгоритм работает с небольшим количеством точек сцены, что позволяет значительно повысить скорость его работы. Классические схемы стереосопоставления ищут соответствия для каждой точки стереокадров, в нашем случае сопоставление производится для точек, соответствующих объектам интереса. Объектами интереса выступают лица людей, обнаруженные на каждом из ракурсов наблюдения. Для детекции лиц используется каскад Хааровских классификаторов [1].

На рисунке 1 представлены два экспериментальных изображения с найденными лицами. На правом кадре черным указаны эпиполярные линии, соответствующие центрам областей на левом кадре.



Садыхов Рауф Хосровович, д.т.н., профессор, зав. каф. электронно-вычислительных машин УО «Белорусский государственный университет информатики и радиоэлектроники».

Ламовский Д.В., аспирант УО «Белорусский государственный университет информатики и радиоэлектроники». Беларусь, БГУИР, 224017, 220013, г. Минск, ул. П. Бровки, 6.

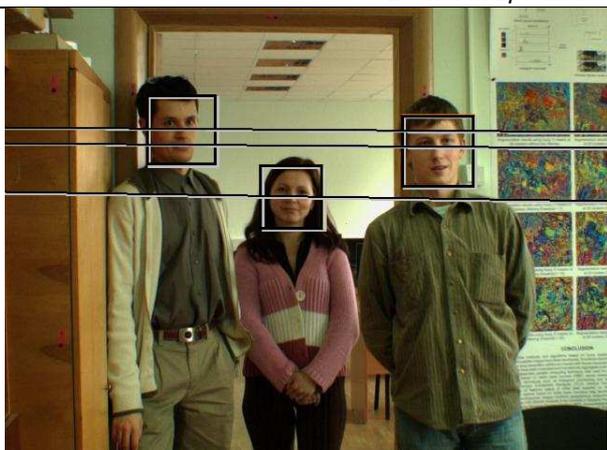


Рис. 1. Левый и правый ракурсы, полученные системой стереозрения. На кадрах представлены найденные лица и соответствующие эпиполярные линии

Алгоритм сопоставления. Для установления соответствия между изображениями лиц на различных ракурсах мы использовали подход, основанный на анализе их геометрических характеристик, положения лица в пространстве, а также на их визуальной корреляции. Лица, полученные на каждом ракурсе, рассматриваются как объекты, характеризующиеся своими геометрическими характеристиками и паттерном. Используя эти характеристики, алгоритм вычисляет степень соответствия объектов с двух ракурсов. Назначение этой процедуры состоит в выборе наиболее вероятных пар и игнорировании лиц, для которых нет возможности найти пару.

Степень соответствия пары лиц определяется произведением набора коэффициентов, каждый из которых соответствует корреляции некоторого параметра:

- $K1$ – определяется точностью положения лица на эпиполярной линии;
- $K2$ – определяется степенью соответствия размеров лиц;
- $K3$ – определяется степенью попадания лица в трехмерную область интереса;
- $K4$ – определяется соответствием гистограмм рассматриваемых областей.

Все вычисления проводятся в координатной системе одного из ракурсов. Из всех возможных пар лиц к рассмотрению принимаются только те, которые удовлетворяют эпиполярному ограничению. То есть те пары, для которых эпиполярная линия, соответствующая центру области лица на одном кадре, проходит через область лица на втором кадре. Это ограничение позволяет сразу отбросить большое количество заведомо ложных пар. Далее вычисляются коэффициенты, зависящие от положения и размера областей. Коэффициент $K1$ рассчитывается как отношение расстояния от центра одной области до эпиполярной линии, соответствующей центру второй области, к минимальному размеру области. Таким образом, указанный коэффициент принимает значение 1 в случае полного попадания центра области лица на эпиполярную линию. Коэффициент $K2$ вычисляется как отношение размера меньшей области к размеру большей. Таким образом, $K2$ всегда меньше или равен единице.

Для вычисления третьего коэффициента необходимо оценить положение лица в пространстве.

Оценка положения лица в пространстве. Из стереометрии известно, что расстояние до объекта сцены можно измерить с помощью триангуляции. На рисунке 1а представлена простая геометрия стереоустановки.

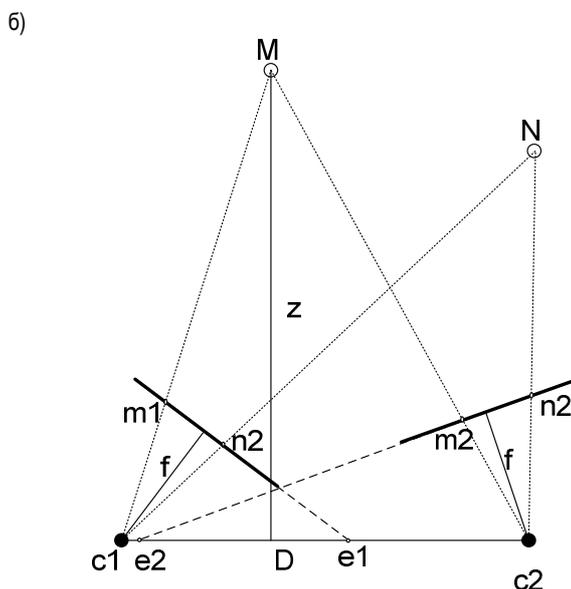
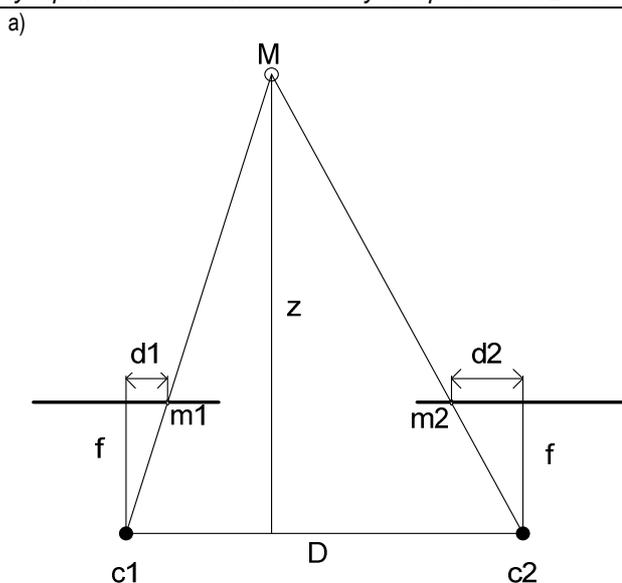


Рис. 2. Геометрия стереоустановки: а – плоскости изображения параллельны оси $c1, c2$; б – свободные положения плоскостей изображения

Здесь $c1$ и $c2$ есть оптические центры камер, M – наблюдаемая точка сцены, $m1$ и $m2$ – проекции этой точки на соответствующие плоскости изображений. Для вычисления расстояния до объекта при такой сцене можно воспользоваться следующим соотношением:

$$\frac{z}{D} = \frac{f}{d1 + d2}, \quad (2)$$

где z – искомое расстояние, D – расстояние между центрами камер, f – фокусное расстояние, $d1 + d2$ – диспаратитет. Диспаратитет, как известно, представляет собой разницу между положениями точки на различных ракурсах.

Реальная стереосистема редко имеет такую идеальную геометрию. На рисунке 1б представлена геометрия стереосистемы с произвольной установкой камер. У такой системы плоскости изображений не параллельны базовой линии (линии, соединяющей оптические центры камер). Для использования выражения (2) при такой конфигурации необходимо применять проективные преобразования к каждой из картинных плоскостей. Это приводит к дополнительным вычислительным затратам. Кроме того, центры изображений перестают быть точками отсчета для вычисления диспаратитета. Так, на рисунке 1б видно, что для двух точек пространства, которые удалены от базовой линии на различные расстояния, эти значения равны.

В таком случае нужны точки, относительно которых можно было бы вычислять диспаратет, который в свою очередь не зависел бы от конфигурации камер. В качестве таких точек мы используем эпиполярные центры, указанные на рис 16 точками e_1 и e_2 . Эпиполярные центры – это точки пересечения всех эпиполярных линий на соответствующем ракурсе, которые могут быть получены с использованием следующих соотношений:

$$\begin{aligned} e_1 F &= 0, \\ e_2 F^T &= 0, \end{aligned} \quad (3)$$

где F – фундаментальная матрица, T – символ транспонирования.

Диспаратет, вычисленный относительно эпиполярных центров не может использоваться для точного определения расстояния до объекта, но позволяет оценить эту величину. Если за нулевую точку ($Z=0$) взять точку, лежащую на базовой линии, то значения диспаратета смогут оказаться слишком большими, особенно, если плоскость изображения имеет небольшой угол относительно базовой линии. Большие значения диспаратета неинформативны. Поэтому за точку отсчета мы выбираем пересечение оптических осей камер. Для этой точки пространства вычисляется базовый диспаратет d_0 . Далее глубину оцениваем по разности диспаратета точки и базового диспаратета. То есть для точек пространства M и N имеем соответственно:

$$dm = dist(m1, e1) + dist(m2, e2), \quad (4)$$

$$dn = dist(n1, e1) + dist(n2, e2),$$

где $dist$ – функция расстояния между точками, dn , dm – абсолютные диспаратеты. При этом оценка глубины будет вычисляться как $dm - d_0$ и $dn - d_0$ соответственно.

Таким образом, коэффициент $K3$ вычисляется как:

$$K3 = 1 - \frac{|dn - d_0|}{\Delta d}, \quad (5)$$

где Δd – разброс диспаратета относительно нулевой точки.

Сопоставление гистограмм. Коэффициенты $K1$ и $K2$ позволяют отбросить львиную долю ложных соответствий, но не дают полной картины соответствия, так как основаны на геометрических параметрах областей, которые могут быть получены со значительными погрешностями. Коэффициент $K3$ также не может в полной мере использоваться для решения вопроса о соответствии. Он может быть использован для определения попадания лица в трехмерную область интереса. Таким образом, для оставшегося небольшого числа возможных пар вычисляется коэффициент соответствия ($K3$), основанный на сопоставлении гистограмм. Данный вид сопоставления был выбран по причине его относительной вычислительной простоты по сравнению со стандартными метриками.

Для построения гистограммы используется не вся область лица, куда могут попасть части фона, волос, одежды, а только центральные области, включающие глаза, нос и рот. Для приведения гистограмм в универсальный вид используется медианный фильтр и нормализация по площади области.

Для сопоставления гистограмм были использованы две метрики: сумма квадратных разностей (SSD) и сумма абсолютных разностей (SAD). Наилучший результат был достигнут при использовании SAD к гистограммам по всем трем каналам цветного изображения. Ошибка гистограммы в этом случае вычислялась как квадрат евклидова расстояния между ошибками по каждой из компонент R, G и B. Для вычисления коэффициента $K4$ использовалось следующее выражение.

$$K4 = 1 - histERR, \quad (6)$$

где $histERR$ – ошибка соответствия гистограммы.

Таким образом, максимизация произведения $K1 * K2 * K4$ дает нам наиболее вероятные пары. Каждое лицо из всего множества может быть включено только в одну пару. То есть наиболее валидная пара, сформированная одним реальным лицом, сразу исключает из рассмотрения все ложные пары, с теми же ракурсами. Ложные пары, состоящие из лиц, для которых нет верных сопоставлений (в силу

необнаружения соответствующего лица на втором ракурсе) отбрасываются по порогу, применяемому к составному коэффициенту. Соответствующий порог был определен эмпирическим путем и равен 0.3.

Результаты. Алгоритм тестировался на изображениях, полученных с помощью стереоустановки, состоящей из двух камер высокого разрешения. Кадры стереопар являются цветными изображениями, имеют разрешение 1024*768 пикселей. В качестве параметров алгоритма было принято, что количество лиц на один стереокадр не будет превышать 5, минимальный размер лиц – 100*100. Количество протестированных стереопар – 200 (табл. 1).

Таблица 1. Результаты исследований

Параметр	Числовое выражение	Процентное выражение
Число лиц на левом ракурсе	336	-
Число лиц на правом ракурсе	372	-
Обнаружено лиц (лев.)	326	97%
Обнаружено лиц (прав.)	366	98%
Всего верных пар	306	-
Верно сопоставлено пар	286	93,5%
Невалидные пары по порогу	4	1,3%
Ложное сопоставление	2	0,65%
Лиц без пары (лев.)	30	13%
Лиц без пары (прав.)	61	

Таким образом, точность сопоставления лиц составила 93.5%. Время на обработку одной пары кадров составило порядка 50 мс, причем основную часть (>90%) этого времени занимает детекция лиц. Таким образом, представленный алгоритм позволяет устанавливать соответствие между ракурсами стереолиц с большой скоростью.

Заключение. Разработанный алгоритм позволяет проводить получение стереолиц с большой вычислительной эффективностью. Алгоритм основан на разреженном стерео сопоставлении объектов, найденных на стереоракурсах. Высокая скорость работы алгоритма достигнута с помощью использования метода оценки глубины, не требующего проективных преобразований картинных областей.

Таким образом, предложенный алгоритм может быть использован в системах видеонаблюдения реального времени как основа для извлечения стереолиц. Стереолица могут служить исходными данными для:

- выбора оптимального ракурса для использования в задаче 2D распознавания;
- построения оптимального ракурса для использования в задаче 2D распознавания;
- построения 3D модели лица для различных задач, в том числе и трехмерного распознавания;

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Viola P. Jones M. J. Robust Real-Time Face Detection // International Journal of Computer Vision – 2004 - vol. 57 - pp. 137-154.
2. A Morphable Model For The Synthesis Of 3D Faces / V. Blanz and T. Vetter // Proceedings of the 26th annual conference on Computer graphics and interactive techniques – 1999 – pp. 187 - 194
3. Zicheng Liu, Zhengyou Zhang, Chuck Jacobs and M. Cohen Rapid Modeling of Animated Faces From Video // The Journal of Visualization and Computer Animation – 2001 – vol. 12 num. 4 - pp. 227-240
4. I. H. Richard In Defense of the Eight-Point Algorithm // IEEE Trans. Pattern Anal. Mach. Intell. – 1997 - vol. 19 – pp. 580-593,.
5. An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration / A. J. Lacey, N. Pinitkarn, and N. A. Thacker // In Proceedings of the British Machine Vision Conference BMVC2000 – 2000.
6. Changhan Park and Joonki Paik Face Recognition Using Optimized 3D Information from Stereo Images // Face Recognition, Book edited by: Kresimir Delac and Mislav Grgic - June 2007 - ISBN 978-3-902613-03-5 - I-Tech, Vienna, Austria - pp.558

7. R. Niese, A. Al-Hamadi and B. Michaelis A Novel Method for 3D Face Detection and Normalization // Journal of Multimedia, Academy Publisher – 2007 – vol. 2, no. 5 - pp. 19-26.

Материал поступил в редакцию 20.09.08

SADYCHOV R.H., LAMOVSKIJ D.V. Definition of the stereopersons in systems of videosupervision

In given clause the algorithm of definition of the persons on the stereostaff of a videosequence is considered. For reception such stereo of videosequences the chambers of the high sanction of systems of videosupervision are used. Received as a result of job of algorithm of the image represent the person of the man who has been removed with different foreshortenings. Such given can be used for reception of the information about structure of the person and construction of its three-dimensional model. For the decision of the specified task the method of definition of a rule of the person on each of the staff of stereopair with use of the cascade Chaar of qualifiers is used effective from the computing point of view [1]. The information on the found persons then incorporates on the basis of algorithm of rarefied stereocomparison. The algorithm is rarefied in the sense that the information on a three-dimensional configuration of a stage is restored not for all of its points, but only for points of interest, that allows to achieve high speed of job of algorithm. As criteria of conformity of pair foreshortenings to one person are considered: performance epipolar of restrictions, conformity of the sizes of areas, rule of received object in 3D of area of interest, and as a degree of correlation histograms considered.

УДК 681.3

Пасичник Роман, Саченко Анатолий

ОБЩАЯ СХЕМА СЕМАНТИЧЕСКОГО WEB-ПОИСКА С ИСПОЛЬЗОВАНИЕМ ПРОБЛЕМНО-ОРИЕНТИРОВАННОЙ ОНТОЛОГИИ

Введение. Проблемы информационного поиска, в частности поисковый шум и потери информации, обычно решают двумя способами - усовершенствованием средств поиска, или улучшением семантики Web. Первый способ, учитывая несложность и приемлемые расходы на реализацию, является более целесообразным [1].

Существующие на сегодняшний день модели информационного поиска базируются на статистике употребления в исследуемых текстах ключевых слов или его семантическом анализе на базе статической онтологии. Первый подход сужает поисковое поле за счет отсутствия учета синонимии слов, а также чрезмерно расширяет его при учете случайного сочетания в различных частях текста отдельных слов, за которыми пользователь предполагал совместную смысловую нагрузку. Снятие этих недостатков возможно за счет использования онтологий, которые должны учитывать не столько конкретные текстовые представления понятий, а весь синонимический ряд таких представлений а также возможные взаимоотношения понятий. Как правило, практическая реализация онтологий предполагает таксономии понятийных классов. Как раз в построении таких таксономий встречаются существенные разночтения. Тут следует вспомнить хотя бы пестрое множество онтологий верхнего уровня - CYC, SUMO, Sowa's, WordNet [2]. Несхожесть ключевых структур их таксономий объясняется разными их предназначениями. Это наводит на мысль о невозможности построения единой древовидной таксономии понятий и необходимости рассмотрения их структуры в виде мультиграфа. Каждый разработчик начинает строить свою онтологию в виде дерева, которое в процессе последующих пополнений вырождается в сетевую структуру.

На взгляд авторов, представление методики разработки онтологий в рамках декларативных онтологий весьма затруднительно. Формирование такой методики удобно осуществлять для их процедурной разновидности в рамках так называемых проблемноориентированных онтологий, последний формализм предложен Р. Пасичником. Как будет показано ниже, первый позволяет провести обоснование принципов первоначального таксонирования онтологий способом, прозрачным для пользователя.

1. Концепция проблемно-ориентированной онтологии. Под онтологией понимаем некоторое формальное описание предметной области, которая широко поддерживается и используется. Общеиспользование опирается на неформальное или формальное соглашение сообщества экспертов относительно конкретного описания предметной области, а формальность позволяет обрабатывать это представление на компьютере. Формализм должен базироваться на

структурировании понятий предметной области, связей между ними, а также их функций, называемых функциями интерпретации. Целью такой формализации можно считать компактное представление знаний о предметной области.

Понятия предметной области можно представить в виде некоторых информационных структур, а в более обобщенном виде – информационных объектов [3, 5]. Для описания таких объектов понятий среди их свойств удобно выделять множество возможных состояний. Для фиксации таких состояний необходимо определение соответствующего множества идентифицирующих их параметров-показателей. Состояния объектов могут изменяться вследствие внешнего воздействия или действия внутренних механизмов объекта. Способы изменения состояний объекта будем называть процедурными знаниями. Под проблемой будем понимать механизм изменения состояния объекта, который важен для исследователя. При этом группировка проблем создает соответствующие структуры, которые можно рассматривать как комплексные проблемы.

С этой точки зрения произвольная предметная область или произвольная структура знаний, объединенная общим подходом к решению некоторой сложной задачи, может рассматриваться как некоторая комплексная проблема. Такой взгляд объясняет наличие существенных различий в построении онтологий верхнего уровня. Исходя из невозможности определить общие приоритеты использования онтологий верхнего уровня, на наш взгляд, следует отказаться от попыток ее формального построения. Более перспективным является подход определения предметных онтологий, имеющих приоритетное значение для отдельных исследователей. Однако базовые понятия этих онтологий должны быть определены через обобщенные понятия, которые выступают метапонятиями разрабатываемой онтологии.

Перечень проблем предметной области считаем ключевым для ее формального описания, поскольку он совмещает множество понятий и подсистем предметной области с множеством методов влияния на нее. Структуризация проблем области приводит и к структуризации знаний о ней. При таком подходе резко уменьшается число степеней свободы структуризации, а, следовательно, растет вероятность согласования мыслей относительно структуры онтологии между разными экспертами, при отсутствии принципиальных расхождений между ними.

Представление методов решения проблем возможно на основании формализации необходимых воздействий в формате близком к описанию функций в программировании. При этом должны быть указаны исследуемый объект, результаты операции, ее название,

Пасичник Роман, Научно-исследовательский институт интеллектуальных компьютерных систем.

Саченко Анатолий, д.т.н., профессор Тернопольского национального экономического университета.

Украина, 46004, г. Тернополь, пл. Победы, 3.