

# ФОРМАЛИЗОВАННОЕ ОПИСАНИЕ СОДЕРЖАНИЯ ИЗОБРАЖЕНИЯ КАК ДАННЫЕ ДЛЯ ГЕНЕРАЦИИ ТЕКСТА

## SEMANTIC ANNOTATION OF AN IMAGE AS THE INPUT FOR NATURAL LANGUAGE GENERATION

*Соколова Е.Г. (eg.sokolova@gmail.ru), Российский государственный гуманитарный университет  
М.В. Болдасов (boldasov@nm.ru), Luxoft*

В данной статье обсуждается представление содержания изображений в виде объектно ориентированных формализованных описаний и их использование в качестве входа для системы генерации текста на естественном языке. Обсуждается структура текстов описаний изображений на естественном языке и средства для создания формализованных представлений содержания изображения, а также типы реально порождаемых моделью текстов.

В последнее десятилетие актуальна задача описания изображений для информационного поиска (ИП) в коллекциях изображений. Нас интересует возможность поиска изображения по его содержанию, т.е. по изображенным объектам и композиции изображения, а не по техническим (цветовая гамма, фактура, наличие определенных статистических характеристик) или библиографическим данным (Соколова, Болдасов, 2006). Мы не можем автоматически получить формализованное представление содержания изображения из его представления в компьютере, т.е. из матрицы пикселей, ввиду сложности проблемы распознавания образов. Но мы можем описывать содержание изображений вручную, учитывая свойства изображенных объектов и особенности зрительного восприятия плоских изображений. Описания изображения на естественном языке (ЕЯ) может рассматриваться как источник для разработки формализованного представления содержания изображения. Описания на ЕЯ непосредственно рассматривались в качестве поискового образа изображения. Однако с ними связано две трудности: 1) они непригодны для многоязыкового поиска; 2) они субъективны. Предлагалось два способа преодоления этих недостатков при ИП: использование информационно-поискового тезауруса (Добров и др., 2000) и перифразирование (Emeljanov et al. 2000). Но ЕЯ описания настолько нерегулярны, что иногда описания одного и того же изображения разными людьми невозможно отождествить с одним изображением, например, в (Добров и др., 2000) приводятся следующие два описания одной и той же фотографии: 1) *Ярко-синее небо, ветки с маленькими листочками, верхняя часть храма белого цвета с зеленым куполом.* 2) *Башня шатрового типа с зеленым шпилем на фоне очень синего неба. Наверху что-то типа герба. Здание окрашено в розовый цвет с белой отделкой.* Чтобы преодолеть эту несходимость содержание изображений или должно описываться в некотором унифицированном виде, или должно быть заменено на формализованное представление содержания изображений. В данной статье мы рассматриваем объектно ориентированное формализованное представление содержания изображений, которое построено на основе изучения ЕЯ описаний, и из которого вновь можно автоматически построить описание содержания изображения на ЕЯ. Конечной целью данного исследования является создание технологии представления содержания изображений и накопление знаний для создания подобных представлений и генерации из них описаний изображений на ЕЯ. Знания впоследствии могут быть использованы для интеллектуализации систем распознавания образов и организации диалога с такими системами. В настоящее время элементы технологии описания содержания изображений и система ГЕЯ используются в преподавании автоматической обработки текстов в РГГУ.

В качестве экспериментального массива изображений были взяты фотографии открытого пространства, предложенные студентами, а затем фотографии с сайта «Коллекция всех фотографий С.М. Прокудина-Горского, восстановленных в цвете»: <http://www.prokudin-gorsky.ru>. С участием студентов 4-го курса РГГУ был создан небольшой экспериментальный корпус ЕЯ описаний этих фотографий и формализованные представления их содержания, которые позволили провести эксперименты по генерации текстов.

### *1. Тексты описания изображений и их организация*

Экспериментальный корпус описаний изображений содержит около ста описаний фотографий открытого пространства на ЕЯ, которые создавалось произвольно, без каких-либо ограничений. Затем создавалось второе

описание, ориентированное на созданную нами для этой цели экспериментальную онтологию понятий, содержащую объекты, представленные на изображениях, их свойства и отношения. Второе описание обычно более лаконично, чем первое, и не содержит информации, полученной не из самого изображения, а из других источников - библиографического описания, догадок автора, например, подчеркнутая часть следующего текста: *На фотографии изображена церковь. На ступеньках церкви сидит мужчина. Может быть он пришел из деревни, которая видна за полем.* Второе описание служило источником для формализованного представления содержания изображения с помощью рассмотренных ниже средств. Такая двухступенчатая методика преследовала две цели: исследовательскую – получить произвольные описания изображений, не ограниченные какими-либо условиями, и получить набросок формализованного представления, а также педагогическую – вызвать у студента потребность в рассуждении о происхождении средств его описания, в частности о том, что существительные имеют прямых референтов в изображении и должны быть представлены в формализованном представлении, а глаголы – нет.

При описании только видимой информации мы можем опираться на наше «наивное» представление о матрице пикселей изображения, учитывая пространственное расположение объектов на плоском изображении и понятие глубины изображения за счет расположения объектов приблизительно на одной горизонтальной прямой, а также вложенных в границы фрагмента изображения, занимаемого другим объектом, или частично скрытых другими объектами. Этот прием позволяет при создании текста описания опираться на объективные свойства содержания изображения. Рассмотрим сначала свойства текстов описаний на ЕЯ.

Каждое предложение текста вводит в повествование один или группу объектов, включая их контекст. Таким объектом иногда становится все изображение, например, *На фотографии мы видим летний день после грозы.* или *На фотографии представлен городской вид.* В остальных случаях предложения вводят конкретные изображенные объекты и группы объектов, например, *На фотографии изображен памятник Петру Великому. Памятник окружен металлической оградой. Перед памятником – цветы.* В текстах представлено два способа группировки объектов - по планам и по группам. Названия групп используются в двух случаях: если это группа однородных объектов, например, *вдалеке видна группа домов;* или если пространство делится на удаляющиеся от наблюдателя плоскости - планы, например: *Это ясный зимний день. [На первом плане мы видим детей, которые играют в снежки на большой лужайке.] (план 1) [На заднем плане – лес.] (план 2).* Группы объектов также образуются вокруг одного выделенного объекта, если образуют с этим объектом один фрагмент изображения, например:

(1) *[На фотографии изображен большой белый дом, вокруг дома сад, окруженный низкой деревянной оградой.] (группа 1) [Перед домом дорога, на которой мы видим две машины. Одна машина зеленая, другая - красная. На другой стороне дороги видны два пешехода – мужчина и женщина.] (группа 2) [Слева от дома детская площадка. На детской площадке двое детей играют в песочнице. Рядом с песочницей – качели и горка.] (группа 3)*

Для создания текста важно указать порядок включения объектов в текст. Теоретически он основывается трех параметрах: 1) семантический вес объекта; 2) стратегия описания; 3) коммуникативная структура текста (предложений текста). Понятие семантического веса вырабатывалось в исследованиях по когнитивной лингвистике при описании визуальных сцен, например, (Кобозева, 1997). Вес складывается из физических, социальных и когнитивных свойств объекта, например, неподвижность, величина, важность для наблюдателя – дом важнее, чем дерево, роль в сцене, например, центральность. Это понятие в западной лингвистике получило название *salience* «выделенность» и широко использовалось в работах по автоматической генерации текста, см., например, докторскую диссертацию «Aspects of salience in NLG»: <http://citeseer.ist.psu.edu/patabhiraman93aspects>. При описании люди обычно группируют объекты опираясь на композицию изображения, т.е. обходя сцену в определенном порядке, в частности: «в порядке убывания весов», «справа налево», «слева направо», «в порядке удаления от наблюдателя» и др. Коммуникативная структура обычно привлекается в системах генерации в те моменты процесса генерации, когда порождающийся текст может иметь более одного продолжения, например, в системе Маккьюин ТЕХТ (Маккьюин, 1989).

## 2. Формализованное представление содержания изображения

На вход универсального генератора должны подаваться четыре вида информации: база знаний (БЗ), цель текста, модель адресата и история дискурса. Последняя создается динамически в ходе генерации текста (Reiter and Dale, 2000), см. также обзор (Болдасов, Соколова, 2005) и (Болдасов, Соколова, 2006). Мы ставим перед собой более скромную цель построения действующего генератора, который эксплицитно решает только часть проблем генерации, но обеспечивает построение текста. Цель и модель адресата при этом «защиты» в модель процесса как это делается во многих практических генераторах. БЗ, однако, представляет связанные с ней труд-

ности в полном объеме. Она делится на две части – входные данные, с одной стороны, представленные формализованным представлением содержания изображения, и ресурсы системы генерации, с другой стороны, т.е. знания различной природы (физические, прагматические, семантические) об изображенных объектах и зрительном восприятии, а также лингвистические средства конкретных языков и знания о них. Знания об изображенных объектах и зрительном восприятии позволяют построить план текста описания, лингвистические средства и знания о них позволяют выразить это содержание средствами конкретного ЕЯ.

Разрабатывая методику создания представлений содержания изображений мы опираемся на объективные свойства изображения. С точки зрения гипотетической системы распознавания образов объекты на изображении представляют собой «пятна», обладающие цветом, формой и фактурой. Именно они должны идентифицироваться ею как конкретные объекты, например, *дом, лес, человек* и т.д. В нашем случае роль распознавателя выполняет человек. К формализованному описанию предъявляются следующие требования: 1) оно должно описывать всю плоскость изображения, т.е. включать все крупные изображенные объекты и значимые объекты фона и перечислять все более мелкие объекты с определенной степенью дробности; 2) оно должно отражать композицию изображения с помощью двух видов отношений: вложение и пространственные отношения. Мы используем бинарные отношения, которые устанавливаются между двумя узлами. Такой способ представления содержания текстов был предложен Н.Н.Леонтьевой (Леонтьева, 1967). В нем предполагалось, что первоначальные отношения (у нас – прагматические) при понимании (у нас - при выражении средствами языка) должны далее интерпретироваться. В нашем случае каждый узел представляет конкретный изображенный объект, группу объектов типа *план* или *группа* или все изображение целиком – *фото*, и может содержать свойства представляемого им простого или сложного объекта, в частности, *цвет, количество, имя* и т.д. Узел *фото* может включать информацию библиографического или географического типа, например, *год съемки, изображенное место* и т.д.

Композиция изображения в целом представлена отношением вложения. Объекты и их группы начиная с самого крупного – фото, вложены друг в друга как матрешки. Вложение – это отношение принадлежности объектов к одному фрагменту изображения или «пятну», которое представлено главным объектом (*фото* или выделенный объект) или групповым узлом типа *план* или *группа*. Объекты изображения в виде узлов и отношения вложения представляются в виде дерева составляющих в формализме XML. Оно составляет первую часть формализованного представления, описывающую композицию (группировку) объектов на изображении.

### 2.1. Пространственные, локативные и пространственно-локативные отношения

Вторую часть формализованного описания составляют пространственные отношения. К ним мы в настоящее время относим пять основных отношений ориентации объектов на изображении с позиции наблюдателя и отношение локализации. Пространственные отношения:

*to-the left/to-the-right, behind/in-front-of, above/under, around/in-the-middle, near/near*

являются обратимыми, т.е. если задано пространственное отношение R (A, B), например, *behind (house, tree)* или *near (house, tree)* – *дом позади дерева* или *дом около дерева*, то существует и обратное ему отношение R' (B, A), например, *in-front-of (tree, house)* – *дерево перед домом* или *near (tree, house)* – *дерево рядом с домом*<sup>1</sup>. Введение обратимых пространственных отношений позволяет применять различные стратегии описания пространства, т.е. вводить объекты в текст в нужном автору текста порядке. Пространственные отношения позволяют передать композицию всего изображения более точно, чем отношения вложения, ориентируя группы объектов друг относительно друга. Пространственные отношения не нуждаются в дополнительной интерпретации, чтобы быть выраженными на ЕЯ, так как они сами являются закрепленными в языке, его семантике, элементами опыта людей, в частности, в семантике локативных предлогов, например, *рядом, около, справа от, позади, below, to the left of* и т.д. Пространственные отношения задают направление от одного объекта к другому, причем либо эти объекты расположены неконтактно, либо это несущественно.

В отличие от пространственных отношений отношение локализации – *localisation (A, B)*, описывает контактное расположение объектов на изображении. Связываемые им объекты обладают определенными взаимными физическими и функциональными свойствами, которые известны людям и являются частью их опыта. Отношение локализации устанавливается между двумя объектами, уже связанными отношением *вложение*, и помещается во вторую часть представления вместе с пространственными отношениями. Оно является интерпретацией отношения вложения, которое, будучи не семантическим, а предметным по своей природе, не может быть

<sup>1</sup> Предметное отношение *near* – симметрично, в отличие от языкового значения предлога, которое предполагает определенное соотношение когнитивных свойств объектов, например, велосипед около стены vs. \*стена около велосипеда.

непосредственно выражено средствами языка. Отношение локализации имеет конкретные средства выражения в языках, в частности, в английском языке оно выражается тремя нейтральными локативными предложениями – *at, on, in*, а в русском языке – двумя: *на, в*. Например, дерево, растущее на лужайке, составляет с ней единый фрагмент изображения и частично вложено в нее. В этом случае отношение вложения может быть дополнительно интерпретировано отношением *localisation (лужайка, дерево) – на лужайке дерево*. Отношение локализации не является двунаправленным, т.е. для отношения *localisation (лужайка, дерево)* нельзя написать обратное или симметричное отношение *\*R'(дерево, лужайка)*. Участники этого отношения сохраняют свои роли: *дерево* – объект, *лужайка* – место (или локус). Чтобы выполнить требование возможности включения в текст понятий в прямом и обратном порядке, название отношения можно переписать в терминах ролей: locus/object. В этом случае получим, как и у пространственных отношений пару отношений: locus/object и object/locus, а порядок вхождения в текст будет выражаться через коммуникативную структуру предложения, например, locus/object – *на лужайке дерево*, object/locus – *дерево на лужайке*. Интерпретация отношения вложения через отношение локализации может следовать из семантики понятий, например, дорога – пространство, автомобиль – артефакт, соответственно, *автомобиль на дороге*, а не наоборот. Но бывают ситуации, когда только изображением определяется, что является локусом, а что объектом, например, *книга на журнале, журнал на книге*.

Особый случай для формализованного представления – это ситуация совмещения отношений локализации и пространственных, например, *в центре дома окно*. В этом случае необходимо введение специальных пространственно-локативных отношений, включающих членение объемлющего объекта на части, например, *в\_верхней(нижней)\_части(A, B), в\_правой(левой)\_части(A, B), в\_центре(A, B), в\_левом(правом)\_нижнем(верхнем)\_углу (A,B)*. Пространственно-локативные отношения так же, как отношение локализации, являются по существу необратимыми.

## 2.2. Атрибутивные отношения и онтология

Относительно всех остальных отношений вложения, не интерпретированных как локализация, мы поступаем единым образом: признаем их атрибутивными, причем объемлющий объект выполняет роль объекта, а вложенный в него – атрибута. Например, *окна* составляют единый фрагмент изображения с *домом*, но они не локализованы *на(в)* нем, а являются его частью. Это отношение реализуется на ЕЯ разными способами: *дом с (тремя) окнами, дом имеет окна, окна дома* и др. Как и отношение локализации, атрибутивное отношение может быть представлено через роли участников: быть объектом, имеющим атрибут, и быть атрибутом объекта. Атрибутивное отношение не эксплицируется в представлении – отношение вложения непосредственно реализуется с учетом семантических свойств объектов. Есть две возможности: объект, находящийся в фокусе, выполняет в отношении роль атрибута или объекта, например, понятия «дом» и «забор»: дом = объект -> *дом с забором*; дом = атрибут -> *забор дома*. Реализация первого полностью зависит от семантики соединяемых понятий, например, *дом с крышей* (целое-часть), *сотрудник по фамилии* (человек-ФИО), *забор из жердей* (целое-(часть=материал)) и др.... Второе отношение обычно реализуется родительным падежом зависимого слова, например, *крыша дома, жерди забора*. Иногда добавляется предлог: *уговища от пальто*. Реализация атрибутивного отношения в системе QGen рассматривалась в (Соколова, 2003). Классификация объектов, необходимая для интерпретации отношения вложения, а также способы реализации пространственных, пространственно-локативных и локативных отношений, должна находиться в ресурсах генератора.

## 3. Ресурсы генератора

Ресурсы представляют постоянные, статические знания генератора о мире, которые используются, а не преобразуются генератором. Для того, чтобы отношение вложения между узлами и группами узлов можно было автоматически интерпретировать, объекты в ресурсах должны быть классифицированы. Наиболее общие классы, полезные для генерации описания открытого пространства:

- Гео-объекты – это части земной поверхности и небо, например, *дорога, море, лужайка, гора, клумба* и т.д. Эти понятия в тексте интерпретируются как фон для остальных объектов. Их обычная семантическая функция – локус, т.е. функция местонахождения для других предметов, например, *дом на горе*.

- Основные объекты – это объекты, имеющие постоянную связь с земной поверхностью, т.е. *дома, деревья, мосты, ограды* и др. Они выделяются своими размерами и постоянством местонахождения. Относительно этих объектов обычно ориентируются подвижные объекты.

- Подвижные объекты – это относительно небольшие объекты, для которых естественна смена местоположения. Это объекты живой природы, например, *человек, животные, птицы*, артефакты, например, *подъемный кран, ящик, бревно*, транспортные средства, например, *самолет, машина*, и мелкие объекты типа *игрушек, сумок*,

цветов (цветы тоже легко меняют местоположение, попадая в букеты или осыпаясь). Эти объекты обычно являются фокусом интереса наблюдателя.

- Части-объектов – это либо объекты, которые обычно функционируют только в составе другого сложного объекта, например, *крыша (дома), рука (человека), хвост (животного), ветка(дерева)*. Некоторые из них могут существовать отдельно от их целого, например, *окно* может стоять около строящегося дома, кирпич может лежать на траве. В этом случае окно находится с домом в пространственном отношении, например, *near/near(house1>window1)*, кирпич локализован на траве *localisation(brick1, grass1)*. В том случае, если часть-объекта уже едина с объектом, отношение вложения не интерпретируется в составе представления, а непосредственно реализуется как объектно-атрибутивное отношение средствами языка в соответствии с онтологическими типами объектов, например, *дом с окнами* (целое-часть), дом из кирпича (объект, материал). Онтологический тип представляет более точную классификацию понятий действительности, приписываемую понятию в онтологии, например, *одежда* как часть видимого образа человека и *окно* как часть дома. В настоящее время нами составлена экспериментальная модель онтологии, содержащая более 100 понятий - объектов, их свойств и отношений.

### 3. Реализация процесса генерации

Для реализации модели мы используем среду DEMLinG (Болдасов, 2003), которая была создана для предшествующих двух моделей генераторов: QGen и генератора выдачи данных из базы данных. Планируемая модель предполагает два главных отличия от предшествующих: 1) в ней требуется более развитый этап планирования, чем в предыдущих; 2) предполагается отделить ресурсы системы от конкретной модели генератора. Мы используем входные данные «модель предметной области», сознательно отказываясь от семантических обобщений, например, Upper Model в среде KPMML <http://www.fb10.uni-bremen.de/anglistik/langpro/webospace/jb/gum/index.htm>. Использование для формулирования входных данных семантических понятий типа actor, material-action и др. ведет к иллюзорности прокламируемой подобными системами независимости от конкретных языков. Система понятий должна быть основана не на языковой иерархии, а на онтологии понятий реального мира, конфигурациям которых ставятся в соответствие непосредственно конструкции языковых средств конкретного ЕЯ. Описание содержания изображения, с нашей точки зрения, является хорошим материалом для такого исследования.

Для создания формализованного представления кроме рассмотренных отношений необходимо составить инвентарь понятий, которыми могут описываться объекты. Мы предполагаем использовать для этой цели слова ЕЯ или искусственные слова, если первых окажется недостаточно. Понятие видимого объекта имеет две составляющих – образ и свойства. Образ объекта можно представить словом ЕЯ, которое только и может его передать<sup>2</sup>. Но кроме этого слова, выбранного из русского, английского или другого языка, мы должны описать свойства, которые отражает его значение, отличающие его от значений других слов, также возможных для называния подобных объектов. Они включают как физические свойства, так и свойства, связанные с системой понятий данного ЕЯ, которые должны быть исследованы для решения нашей задачи.

#### 3.1. Свойство визуальности в лексическом описании слова

Подобно новым свойствам в лексических описаниях слов, например, коммуникативные свойства (Новый объяснительный словарь синонимов, Предисловие, 2000, с. IX), мы предлагаем ввести признак «визуальность». Этот признак имеют слова, обозначающие объекты реальности, которые могут вводить понятие в текст в интродуктивных предложениях без зависимых членов, обозначающих не визуальные свойства например, *Это дом; На фотографии мы видим высокое здание*. Слова *здание* и *дом* имеют признак визуальности. Многие слова, которыми мы можем обозначить в тексте тот же дом, признака визуальности не имеют, например, *общежитие, жилье, даже дача*. Последнее слово сохраняет семантику процесса – «участок, данный в пользование или собственность», поэтому в интродуктивном предложении появляется с уточнением собственника (данный кому), например, *Это дача Сталина*, но \**На фотографии мы видим дачу*. Признак визуальности значим в системе употребления языка и не зависит от наличия или отсутствия у носителей языка представления о типичном образе объекта, например, у нас есть представление о типичном шестисоточном дачном домике, или пригородной даче начала 20 в., но это не сообщает слову *дача* признака визуальности. Одна из фотографий, изображающая такой дом, в нашем корпусе описывается так: *«Позади справа на траве помещается желтое деревянное здание. Вокруг дома размещены клумба с цветами, фонарный столб, небольшие кусты. На заднем плане различим забор, рядом*

<sup>2</sup> Образ может быть представлен также картинкой, представляющей типичный объект данного типа, которая может использоваться системой распознавания образов.

с которым - часть красного легкового автомобиля, за ними – густой лес.» Основательный дачный дом с типичными атрибутами дачи – жилое загородное строение, клумба, забор, лес, машина, представлен даже не словом дом, а словом здание. В нашем корпусе и в Национальном корпусе русского языка слово дача без невизуальных уточняющих членов в функции представления изображенного или видимого объекта, нам не встретилась.

### 3.2. Проблемы представления предметных понятий в онтологии по обслуживанию изображений

Проблема создания инвентаря понятий состоит в том, что в современных словарях в значениях слов не выделяются составляющие визуальной природы. Рассмотрим, например, статью ДОМ1 в Новом объяснительном словаре русского языка (Новый..., 1997), по содержанию наиболее близком для решения нашей задачи. В нем приводится синонимический ряд слова дом из пяти членов, про которые указывается, что они различаются свойствами пяти типов: «размер», «материал», «функциональность», «изолированность», «время». Их можно представить следующим списком (значение признака выделено жирным шрифтом, признак с тильдой - факультативный):

- дом (размер, материал, **~жилой**, **~изолированный**, время);
- здание (**большое**, **~каменное**, функциональность, **~изолированное**, время);
- строение (размер, материал, функциональность, **~изолированное**, время);
- постройка (**небольшая**, **~деревянная**, **~хозяйственная**, **~изолированная**, время);
- корпус (размер, материал, функциональность, **~один-из-ряда**, **современный**).

Любой из вышеприведенных признаков в конкретной ситуации может быть неверным, например, *Фабрика «Красный Октябрь» – это третий дом от угла* (т.е. не жилой), *Старое маленькое деревянное здание церкви Вознесения сгорело еще в 15 веке* (т.е. не большое, не каменное). *«Завар Кили — это место, расположенное вблизи пакистанской границы, в котором находятся система пещер и комплекс строений, используемые “Аль-Каидой” в качестве лагеря и центра подготовки. [Александр Юрьев. Операция “Несгибаемая свобода” // “Солдат удачи”, 2003]»*, (т.е. не изолированные). *«Простите, — вежливо спросил я его, — знаете ли вы, что вон та большая красная постройка — это не трамвайное депо, а фактически памятник первым метростроевцам? [Василий Гулин, Константин Михайлов. Царское метро // “Столица”, 1997.03.18]»* (т.е. не небольшая). *«Все они были совершенно пусты, только в конце второго помещения выступала толстым цилиндром какая-то постройка из плотно сложенных серых камней. [Иван Ефремов. Обсерватория Нур-и-Дешт (1944)]»* (т.е. не деревянная, в корпусе встречаются постройки из кирпича, бетона). Слово корпус обозначает не обязательно множественное строение, например, *Старое здание завода снесли и построили новый корпус из стекла и бетона*. Признак «современный» следует уточнить, поскольку в Национальном корпусе русского языка есть примеры прошлого и даже позапрошлого, 19 века: *Занимал он несколько комнат в среднем корпусе, из которых одна была обращена в контору. [И.А. Гончаров. Май Месяц в Петербурге (1891)]*. Из пяти вышеперечисленных свойств только два визуально определимы – размер и материал. Таким образом, описанные признаки без дополнительного исследования непригодны к использованию в формализованной модели, занимающейся описанием содержания изображений. Требуется также изучение роли образа и свойств объекта при использовании слов в тексте. Так, например, игрушечный домик явно не обладает признаками размера понятия дом, но является воплощением его образа, а домик из листьев, сделанный ребенком для бабочки, не имеет образа дома, но соответствует другой части его значения – жилой.

В среде DEMLinG есть средства представления словаря, включающие, в частности, возможность указать структурный контекст, в котором данное понятие имеет другую лексическую реализацию. Однако возможность представления в нем универсального ресурса требует изучения.

### 3.3. Источники глаголов

Генерация описаний изображений из формализованных представлений требует кроме рассмотренной выше задачи создания этих представлений и задания языка для описания содержания изображения решения второй основной задачи - разработки средств и правил выражения этих понятий на ЕЯ. Кроме выбора квазисинонима или слова, находящегося с данным в других отношениях, например, гиперонима, для выражения объекта в тексте, необходимо описать способы появления в тексте глаголов, без которых текст не может быть создан. В формализованном представлении понятия процессов отсутствуют. Мы можем их вывести из отношения **локализации** и **поз** живых существ - людей и животных, например, *человек (собака) сидит (бежит, идет, лежит)*, отношения **локализации** объектов определенных типов, например, *бревна лежат, столб стоит* и т.д., а также транспортных средств и гео-объектов, например, *машина на дороге (едет), лодка на реке (плывет)*. Таким образом, в языковой реализации пространственных отношений участвуют глаголы движения, местонахождения,

например, *находится, располагается*, глаголы бытия, например, *есть*, глаголы зрительного восприятия, например, *виден, мы видим, можно заметить, виднеется*, глаголы, включающие в себя параметры объекта, например, *вьется, возвышается, ведет, выступает, занимает, делит пополам* и др. Построение правил использования этих глаголов в генераторе текстов описания пространства также задача ближайших исследований.

#### 4. Экспериментальные стратегии планирования

Входное представление состоит из двух частей – часть «узлы», организованную отношением вложения, и часть «отношения», которая содержит пространственные отношения и отношение локализации между объектами. В настоящее время мы используем две экспериментальные стратегии построения текста: стратегия, основанная на композиции изображения (часть «узлы») и стратегия, основанная на части «отношения». В первом случае процесс построения текста следует структуре узлов, последовательно их перебирая. Каждое построенное предложение вводит новый узел. Получается последовательное перечисление изображенных объектов и текст получается довольно гладким, но монотонным. Во втором случае мы основываемся на отношениях (эта часть представления может быть пустой). Отношения не представляют последовательного и полного перебора узлов, и нельзя указать разумного порядка ввода отношений в текст без учета композиции изображения, представленной в первой части. В результате порождается не очень связный текст. В этом случае также возникает проблема номинации узлов. В номинацию узла, имеющего вложенные узлы, не всегда можно включать лексикализации всех этих узлов. Повторение подробных номинаций делает текст многословным и неуклюжим. Для создания интегральной стратегии, которая позволила бы строить хорошие тексты описания изображений, необходимо совместить эти две стратегии – следовать композиции описания (узлы) и одновременно вводить их отношения с другими узлами из части (отношения). Кроме того, хотелось бы научиться строить текст по запросу пользователя.

#### 5. Заключение

В статье представлены результаты предварительного исследования по созданию генератора текстов описаний изображений. Описаны средства и организация формализованного описания изображения, которое может служить входом для такого генератора. Сформулированы проблемы создания ресурсов генератора, в частности, в виде онтологии понятий. Продолжение исследований включает создание онтологии для понятий формализованных представлений содержания изображений; формализация способов их выражения на ЕЯ; изучение структуры текстов описаний и создание стратегий построения описаний различного типа; развитие среды DEMLinG для расширения ее возможностей по построению многоязыковых генераторов в узких предметных областях.

#### Список литературы

1. Болдасов М.В. Парадигмы генерации ЕЯ текстов в инструментальной среде DEMLinG // Труды международной конференции по компьютерной лингвистике и интеллектуальным технологиям ДИАЛОГ'2003. Протвино: 2003. С. 66-75.
2. Болдасов М.В., Соколова Е.Г. Генерация текстов на естественном языке - состояние вопроса и прикладные системы // НТИ, Серия 2, N10, 2005, с.12-22.
3. Болдасов М.В., Соколова Е.Г. Генерация текстов на естественном языке - теории, методы, технологии // НТИ, Серия 2, N7, 2006. С. 1-15.
4. Добров Б.В., Сидоров А.В., Юдина Т.В., Лукашевич Н.В., Салий А.Д. Постановка эксперимента по индексированию корпуса изображений с помощью специализированного тезауруса // Вторая всероссийская научная конференция «Электронные библиотеки: перспективные методы и технологии, электронные коллекции», Протвино, 2000.
5. Кобозева И.М. Как мы описываем пространство, которое видим: композиционные стратегии // Труды Международного семинара Диалог'97 по компьютерной лингвистике и ее приложениям. Ясная Поляна, 10-15 июня. М. 1997. С. 132-136.
6. Леонтьева Н.Н. Об одном способе представления смысла текста // ИПС и автоматизированная обработка научно-технической информации, Т. II, - М., 1967.
7. Макьюин К. Дискурсивные стратегии для синтеза текста на естественном языке // НЗЛ. Вып. XXIV. М., 1989. С. 311-356.
8. Новый объяснительный словарь синонимов русского языка. Вып. 1, М.: Школа «Языки русской культуры», 1997.

9. Соколова Е.Г. Семантические свойства имен для генерации запросов к базе данных // *Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции Диалог'2003*, Протвино, 2003. С. 640-647.
10. Соколова Е.Г. Болдасов М.В. Принципы построения семантических аннотаций содержания изображений // *Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции Диалог'2006*, Бекасово, 31 мая – 4 июня 2006. С. 474-480.
11. Boldasov M.V., Sokolova E.G. QGen – Generation module for the register restricted InBASE system // *Gelbukh (ed.). Computational linguistics and intelligent text processing, Lecture notes in computer science, N 2588, Springer-Verlag, CICLing 2003, Mexico city, Mexico, 2003. P. 465-476.*
12. Emeljanov G.V., Krechetova T.V., Kurashova E.P. Tree grammars in the problems of searching for images by their verbal descriptions // *Pattern recognition and image analysis*. – 2000. Vol. 10, N4. P.520-526.