

МЕТОДИКА РАЗРАБОТКИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ЗАДАЧАХ КОМПЬЮТЕРНОЙ ОПТИКИ

И.И. Бикмулина, И.А. Барков

Казанский национальный исследовательский технический университет имени Туполева А.Н., Казань, Россия

Рассматривается методика описания предметной области для автоматизации задач компьютерной оптики, которая позволяет передать этап описания предметной области эксперту в оптике.

Ключевые слова: компьютерная оптика, информационные процессы и технологии, проектирование систем.

Введение

Многие задачи компьютерной оптики предполагают (требуют) для их решения использование информационных технологий. Современные тенденции создания информационных технологий подразумевают передачу предварительного описания предметной области специалисту по данной предметной области. Существующие средства описания предметной области (язык UML, диаграммы IDEF, диаграммы бизнес процессов [1] и др.) вызывают затруднения у специалистов предметных областей, в том числе, у специалистов компьютерной оптики: необходимость изучения, трудности корректного использования, дополнительные трудозатраты и т.д. Поэтому возникает необходимость создания таких изобразительных средств, которые могли бы снизить перечисленные трудности. Во-первых, уровень этих изобразительных средств должен быть достаточен для полноценного описания предметной области, во-вторых, средства описания предметной области должны быть интуитивно понятны специалисту предметной области, в-третьих изобразительные средства должны обладать возможностью применения инструментов автоматизации разработки информационных технологий.

В работе [2] предложен язык описания предметной области, уровень которого, по мнению авторов, выше уровня языка UML, более понятен специалисту и позволяет выполнить автоматический синтез UML моделей. На данном этапе из всего множества диаграмм UML выбраны диаграммы классов и для них созданы соответствующие программные средства синтеза диаграмм. Как будет видно из приведенных примеров, часть элементов диаграммы классов не декларируется явным образом. Например, структурные отношения между классами формируются автоматически в процессе синтеза диаграммы классов. Также автоматически вводятся области видимости (`public`, `private`, `protected`) полей, методов, свойств. Благодаря этим введениям язык описания предметной области становится проще, а специалист предметной области освобожден от продумывания реализации диаграммы классов.

В данной публикации на примере произвольно выбранной статьи из журнала «Компьютерная оптика» [3] рассматривается методика описания предметной области.

1. Методика описания предметной области

В [3] рассматривается задача создания технологий и методов бесконтактной идентификации загрязняющих веществ и производств, выбросы которых не удается определить в непосредственном контакте. При этом используются фурье - спектрометры на базе интерферометров.

Принимая указанную статью за первоначальное описание экспертов компьютерной оптики в области спектрометрии, рассматриваем язык и методику формирования описания предметной области, пригодного для автоматического синтеза соответствующей информационной технологии[4].

Основной единицей языка является понятие. Понятие - это свойство предметной области, рассматриваемое как некоторое высказывание, непосредственно связанное с понятием смысла. Смысл, содержание понятия будем выражать с помощью отношений и высказываний. Рассматриваются структурные, атрибутивные отношения, которые определяют соответствующие свойства понятия. С помощью высказываний задаются произвольные свойства понятия, выражаемые с использованием компонентов отношений. Более подробно структура понятия будет рассмотрена ниже.

Дальнейшее изложение строится следующим образом: из [3] выбирается цитата, содержание которой требуется автоматизировать. Выбранная цитата трансформируется в понятия, которые, в свою очередь, позволяет применить средства автоматического синтеза информационной технологии (сам синтез здесь не рассматривается).

Цитата 1 из [3]: «Обозначим через r^j коэффициент корреляции спектра \vec{x} с веществом из базы данных под номером $j = 1, \dots, M$, где M – число веществ в базе данных. Будем считать, что база включает в себя вещество со спектром $\vec{\sigma}^*$, и пусть его номер j^* ».

В данном предложении речь идет о том, что имеется база данных, в которой хранится информация о веществах и их спектрах. Поэтому необходимо ввести понятие «база данных», содержанием этого понятия является: количество веществ, спектр $\vec{\sigma}^*$ для каждого вещества с номером j^* . Определим понятие «база данных»:

Понятие база_данных

Атрибуты Имя_базы_данных: {строка}.

Количество_веществ: {целое}.

Число_точек_в_спектре: {целое}.

Функция Получить_доступ_к_базе_данных (Имя_базы_данных) пусто.

Функция Получить_доступ_к_веществу (Номер_вещества : целое) пусто.

Функция Вычислить_количество_веществ (Имя_базы_данных) Количество_веществ.

Функция Вычислить_число_точек_в_спектре (Имя_базы_данных) Число_точек_в_спектре.

Свойство Номер_вещества из Получить_доступ_к_веществу \leq Количество_веществ (Имя_базы_данных).

В данном примере вводится понятие «база данных», описание понятия начинается с «**Понятие** база_данных». Введенное понятие характеризует множество баз данных, которые могут быть построены с использованием данного описания. Сказанное в равной мере относится и к другим понятиям. В разделе «**Атрибуты**» вводится атрибутивное отношение вида <представитель базы данных, элемент раздела «атрибуты»>. В нашем слу-

чае, если мы построим конкретный объект «База_экспериментальных_данных» с использованием введенного понятия «база_данных», то будем иметь отношение <База_экспериментальных_данных, Имя_базы_данных>, смысл данного отношения очевиден: «База_экспериментальных_данных» характеризуется именем (имеет атрибут) «Имя_базы_данных». Приведенный пример показывает правило: имена понятий начинаются с малой буквы, а имена конкретных представителей понятия (объектов) начинаются с заглавной буквы. В свою очередь, атрибут «Имя_базы_данных» является представителем понятия «строка». Как обычно, область действия атрибута совпадает с содержимым соответствующего понятия.

После определения атрибутов в нашем примере следует описание функций. Описание каждой функции начинается служебным словом «Функция», в рассматриваемом примере приводится описание трех функций. После слова «Функция» стоит наименование функции, далее, в скобках, идет список параметров функции. Заканчивается описание указанием возвращаемого функцией вычисленного значения. В нашем примере в двух описаниях функций возвращаемое значение нас не интересует, что выражается указанием «*пусто*». В третьем случае функция «Количество_веществ» должна вернуть целую константу, значением которой, как предполагается, будет количество веществ в базе данных. На данном уровне рассмотрения (имеется в виду задача построения диаграммы классов) нас интересует структура (сигнатура) функции. В первом и третьем определениях функции параметр «Имя_базы_данных» указывает на соответствующий атрибут определяемого понятия. Во втором определении параметр «Номер_вещества : целое» означает подстановку некоторого целого числа, названного «Номер_вещества».

В описании понятия «база_данных» присутствует свойство, начинающееся со служебного слова «Свойство». В нашем случае свойство представляет собой ограничение, суть которого заключается в том, что, если происходит вызов функции, то вещество с номером «Номер_вещества» должно находиться в базе данных. Здесь используются следующие изобразительные приемы. Выражение «Номер_вещества *из* Получить_доступ_к_веществу» означает, в описании понятия имеется имя «Получить_доступ_к_веществу», в нашем случае это имя функции. Служебное слово «*из*» предписывает найти имя «Номер_вещества» в описании соответствующей функции.

Предполагаем, что в информационной технологии сведения о спектрах веществ будут переноситься из базы данных в некоторый массив:

Понятие массив_эталлоных_спектров_веществ

Атрибуты Массив_эталлоных_спектров_веществ: {вектор[база_данных ::Количество_веществ]из вектор [база_данных ::Число_точек_в_спектре] из вещественное}.

Функция Загрузить_массив_эталлоных_спектров_веществ (Массив_эталлоных_спектров_веществ, Имя_базы_данных) Число_точек_в_спектре.

Свойство Число_точек_в_экспериментальном_спектре > 0.

Цитата 2 из[3]:« Предположим, что исследуемый вектор зашумлен белым аддитивным шумом, имеющим нормальное распределение, тогда итоговый спектр будет иметь такие же характеристики шума из-за свойств преобразования Фурье. То есть исследуемый спектр может быть представлен в виде:

$$\bar{x} = \bar{\tau}^* + \bar{\xi}, \quad \bar{y} = \bar{\tau}^*, \quad (2)$$

где $\bar{\tau}^*$ – вектор эталонного спектра, $\bar{\xi}$ – шумовой вектор, каждая компонента которого имеет нормальное распределение с известными характеристиками $N(0, \sigma_{\xi}^2)$.

Понятие спектр_исследуемого_вещества

Атрибуты Спектр_исследуемого_вещества: {вектор [Число_точек_в_спектре] из вещественное}.

Функция Заполнить_массив (Спектр_исследуемого_вещества) пусто.

Понятие шумовой_вектор

Атрибуты Шумовой_вектор : { вектор[база_данных ::Количество_веществ]из вектор [база_данных ::Число_точек_в_спектре]}.

Функция Вычислить_шумовой_вектор(Массив_эталонных_спектров_веществ, Спектр_исследуемого_вещества) Шумовой_вектор

Цитата 3 из[3]: «Обозначим через r^j коэффициент корреляции спектра \bar{x} с веществом из базы данных под номером $j=1, \dots, M$, где M – число веществ в базе данных. Будем считать, что база включает в себя вещество со спектром $\bar{\tau}^*$, и пусть его номер j^* .

Получим явное выражение для коэффициента корреляции в случае, когда спектр \bar{x} сравнивается со спектром $\bar{\tau}^*$. Получим:

$$r^{j^*} = \frac{\hat{\sigma}_{\tau^*} + \sigma_{\xi} r_{\xi}^{j^*}}{\sqrt{\hat{\sigma}_{\tau^*}^2 + \delta^2 + 2\hat{\sigma}_{\tau^*} \sigma_{\xi} r_{\xi}^{j^*}}} \quad (3)$$

где $\hat{\sigma}_{\tau^*}^2 = \frac{\|\bar{\tau}^* - \bar{\tau}\|^2}{N}$ – среднеквадратичное отклонение спектра $\bar{\tau}^*$, характеризующее его интенсивность,

$$r_{\xi}^{j^*} = \frac{\bar{\xi}^T (\bar{\tau}^* - \bar{\tau})}{N \sigma_{\xi} \hat{\sigma}_{\tau^*}} \quad (4)$$

– коэффициент корреляции чистого шума с эталонным спектром $\bar{\tau}^*$,

$$\delta^2 = \frac{\|\bar{\xi} - \bar{\xi}\|^2}{N} \quad (5) -$$

среднеквадратичное отклонение шума,

N – число точек в экспериментальном спектре.

Принимаем $\delta^2 = \sigma_{\xi}^2$.

Понятие среднее_значение_вектора_эталонного_спектра

Атрибуты Среднее_значение_вектора_эталонного_спектра: вещественное

Функция Вычислить_среднее_значение_вектора_эталонного_спектра (Массив_эталонных_спектров_веществ, база_данных ::Число_точек_в_спектре, база_данных ::Количество_веществ) Среднее_значение_вектора_эталонного_спектра.

Свойство

база_данных::Число_точек_в_спектре > 0.

Понятие среднеквадратичное_отклонение_эталонного_спектра

Атрибуты Среднеквадратичное_отклонение: {вектор[база_данных ::Количество_веществ] из вещественное}.

Функция Вычислить_среднеквадратичное_отклонение (Массив_эталонных_спектров_веществ, Среднее_значение_вектора_эталонного_спектра, база_данных::Число_точек_в_спектре) Среднеквадратичное_отклонение.

Свойство база_данных::Число_точек_в_спектре > 0, база_данных::Количество_веществ.

Понятие среднее_значение_шумового_вектора

Атрибуты Среднее_значение_шумового_вектора: вещественное

Функция Вычислить_среднее_значение_шумового_вектора (Шумовой_вектор, база_данных::Количество_веществ) Среднее_значение_шумового_вектора.

Свойство

база_данных::Число_точек_в_спектре > 0.

Понятие среднеквадратичное_отклонение_шума

Атрибуты Среднеквадратичное_отклонение: {вектор[база_данных::Количество_веществ] из вещественное}.

Функция Вычислить_среднеквадратичное_отклонение (Шумовой_вектор, Среднее_значение_шумового_вектора, база_данных::Число_точек_в_спектре) Среднеквадратичное_отклонение.

Свойство база_данных::Число_точек_в_спектре > 0, база_данных::Количество_веществ.

В понятии коэффициента корреляции эталона и шума в случае, когда спектр \bar{x} сравнивается со спектром $\bar{\tau}^*$ параметр «коэффициент корреляции чистого шума с эталонным спектром $\bar{\tau}^*$ » будет атрибутом.

Понятие массив_коэффициентов_корреляции_эталона_и_шума

Атрибуты Массив_коэффициентов_корреляции_чистого_шума_с_эталонным_спектром: { вектор[база_данных::Количество_веществ]из вектор [база_данных::Число_точек_в_спектре] из вещественное},

Массив_коэффициентов_корреляции_эталона_и_шума: { вектор[база_данных::Количество_веществ]из вектор [база_данных::Число_точек_в_спектре] из вещественное}.

Функция Вычислить_массив_коэффициентов_корреляции_чистого_шума_с_эталонным_спектром (Массив_эталонных_спектров_веществ, Среднее_значение_вектора_эталонного_спектра, база_данных::Число_точек_в_спектре, Шумовой_вектор, среднеквадратичное_отклонение_эталонного_спектра::Среднеквадратичное_отклонение, среднеквадратичное_отклонение_шума::Среднеквадратичное_отклонение) Массив_коэффициентов_корреляции_чистого_шума_с_эталонным_спектром.

Функция Вычислить_массив_коэффициентов_корреляции_эталона_и_шума (среднеквадратичное_отклонение_эталонного_спектра::Среднеквадратичное_отклонение, среднеквадратичное_отклонение_шума::Среднеквадратичное_отклонение, Массив_коэффициентов_корреляции_чистого_шума_с_эталонным_спектром, база_данных::Количество_веществ) Массив_коэффициентов_корреляции_эталона_и_шума.

Свойство

база_данных::Число_точек_в_спектре > 0,

база_данных::Количество_веществ.

Цитата 4 из[3]: «Введем величину Φ_j :

$$\Phi_j = r_{\xi}^{j*} - r_{\xi}^j = \frac{\xi^T (\bar{\tau}^* - \bar{\tau}^*)}{N\sigma_{\xi}\bar{\sigma}_{\tau^*}} - \frac{\xi^T (\bar{\tau}^j - \bar{\tau}^j)}{N\sigma_{\xi}\bar{\sigma}_{\tau^j}} \rangle.$$

Понятие величина Φ

Атрибуты Велечина Φ : {вещественне}

Функция Вычислить_ величину Φ (Шумовой_ вектор, Мас-сив_ эталлоных_ спектров_ веществ, Среднее_ значение_ вектра_ эталонного_ спектра, среднеквадратичное_ отклонение_ эталонного_ спектра.: Среднеквадратичное_ откло-нение, среднеквадратичное_ отклонение_ шума.: Среднеквадратичное_ отклонение, ба-за_ данных.: Число_ точек_ в_ спектре) Велечина Φ .

Цитата 5 из[3]: «Вероятность правильного обнаружения можно представить как функ-цию только одного случайного вектора

$$\bar{\Omega} = \{\Phi_1, \Phi_2, \dots, r_{\xi}^{j*}, \dots, \Phi_M\},$$

имеющего длину M (число веществ в базе данных), и на позиции j^* стоит величина r_{ξ}^{j*} ».

Понятие случайный_ вектор

Атрибуты Вектор: {вектор[база_ данных ::Количество_ веществ] из вещественное}.

Функция Вычислить_ плотность_ вероятности (велечина Φ) Вектор.

Свойство Количество_ веществ > 0 .

Цитата 6 из[3]: «Получим итоговую ковариационную матрицу:

$$\Sigma = \frac{1}{N} \begin{pmatrix} 2(1-r^{1j*}) & \dots & 1-r^{1j*} & \dots \\ \vdots & \ddots & \vdots & \vdots \\ 1-r^{1j*} & \dots & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \rangle.$$

Понятие итоговая_ ковариационная_ матрица

Атрибуты Итоговая_ ковариационная_ матрица: {вектор[база_ данных ::Количество_ веществ]из вектор [база_ данных ::Число_ точек_ в_ спектре] из веще- ственное}.

Функция Вычислить_ итоговую_ ковариационную_ матрицу (Массив_ коэффициен- тов_ корреляции_ эталона_ и_ шума) Итоговая_ ковариационная_ матрица.

Цитата 7 из[3]: «Совместная плотность вероятности для вектора $\bar{\Omega}$ будет иметь вид

$$w(\Omega) = \frac{1}{\sqrt{(2\pi)^M |\Sigma|}} \exp\left(-\frac{1}{2} \Omega^T \Sigma^{-1} \Omega\right) \rangle \quad (15)$$

Понятие плотность_ вероятности_ для_ случайного_ вектора

Атрибуты Плотность_ вероятности_ для_ случайного_ вектора: вещественное.

Функция Вычислить_ плотность_ вероятности (база_ данных ::Количество_ веществ, Случайный_ вектор, Итоговая_ ковариационная_ матрица) Плот- ность_ вероятности_ для_ случайного_ вектора.

Цитата 8 из[3]: «По формуле (16) находится вероятность правильного обнаружения. В итоге вероятность (11) может быть найдена как интеграл от (15) по области допустимых значений:

$$P_{correct} = \int_{\bar{A}} w(\Omega) d\Omega \quad (16)$$

где G – область, ограниченная G_{up} , G_{down} и неравенством (12)».

Понятие оценка_правильности_обнаружения

Атрибуты Интервал_ограничения: { вектор[база_данных::Количество_веществ] из вещественное }.

Вероятность_правильного_обнаружения: {вещественное}.

Функция Вычислить_интервал_ограничения (Массив_эталлоновых_спектров_веществ, Среднее_значение_вектора_эталонного_спектра, база_данных::Число_точек_в_экспериментальном_спектре Шумовой_вектор, среднееквадратичное_отклонение_эталонного_спектра:: Среднееквадратичное_отклонение, среднееквадратичное_отклонение_шума:: Среднееквадратичное_отклонение},) Интервал_ограничения.

Функция Вычислить_вероятность_правильного_обнаружения (Плотность_вероятности_для_случайного_вектора, Случайный_вектор, Интервал_ограничения, Интервал_ограничения) Вероятность_правильного_обнаружения.

Заключение

На основе данной методики получаем описание рассматриваемой предметной области непосредственно от эксперта в области оптики. Что позволяет в дальнейшем в исследовании синтезировать структурные модели, на основе которых будет разрабатываться программное обеспечение для информационных технологий компьютерной оптики.

При использовании предложенной методики появляется возможность применять средства автоматизации проектирования информационных технологий.

Литература

1. Технологии разработки программного обеспечения/ С. Орлов.– М.: Физматлит, 2003. – 688 с.
2. Бикмуллина, И.И. Семантическое моделирование учебных задач в электронных интеллектуальных образовательных средах / И.И. Бикмуллина, И.А. Барков // Вестник КГТУ им. А.Н.Туполева. –2012. – № 2.
3. Морозов, А.Н. Статистическая оценка вероятности правильного обнаружения веществ в ИК Фурье-спектрометрии / А.Н. Морозов, И.В. Кочиков, А.В. Новгородская, А.А. Сологуб, И.Л. Фуфурин // Компьютерная оптика. – 2015. – Т.39№ 4. – С. 614-621. – ISSN 0134-2452.
4. Бикмуллина, И.И. Разработка методики автоматизированного синтеза информационных систем на основе семантических отношений предметной области / И.И. Бикмуллина, И.А. Барков, А.П. Кирпичников // Вестник технологического Казанского университета. –2015. – Т.18. №7. – С. 236-243.