

СЕКЦІЯ 26. СТРАХОВИЙ ФОНД ДОКУМЕНТАЦІЇ: АКТУАЛЬНІ ПРОБЛЕМИ ТА МЕТОДИ ОБРОБКИ І ЗБЕРІГАННЯ ІНФОРМАЦІЇ

ПОШУК ТА ВІДОБРАЖЕННЯ ІНФОРМАЦІЇ НА РАСТРОВИХ ЗОБРАЖЕННЯХ ТЕКСТОВИХ ДОКУМЕНТІВ

Баранцев А. Ю.

Науково-дослідний, проектно-конструкторський та технологічний інститут мікрографії, м. Харків

Впровадження інформаційних технологій в усі сфери діяльності сучасного суспільства ставить питання розробки простих та ефективних засобів роботи з електронними документами. Особливо це стосується растрових зображень, характерною рисою яких є необхідність використання складних процедур для здійснення пошуку та виділення необхідної інформації.

Растрові зображення документів використовуються у разі, коли до електронного документа висувається вимога максимального наближення зовнішнього вигляду до оригіналу. Наприклад, до документів, які є об'єктами культурної спадщини. Для забезпечення можливості контекстного пошуку та наочного відображення його результатів пропонується використання спеціального механізму, який полягає в попередній побудові об'єктної моделі тексту на зображенні – текстової мапи, яка забезпечить координатну прив'язку текстових символів до відповідного зображення. Автором був розроблений редактор-маркер зображення E-Colorer, демонстраційна сторінка якого знаходиться за адресою <http://micrography.gov.ua/ecodemo>. У якості вихідного зображення використовується файл у форматі JPEG. У якості файлу текстової мапи використовується файл у форматі XML. Контекстний пошук виконується у файлі текстової мапи. Результатом пошуку є координати регіонів, які потрібно промаркувати. Маркування зображення зводиться до заливання частин зображення, які відповідають знайденим словам, символам тощо, певним кольором. Заливання кольором стає можливим завдяки тому, що між яскравістю текстового символу і фону існує певний контраст.

Ручна побудова текстових мап – трудомісткий процес. Автором був розроблений спеціалізований редактор мап E-Mapedit, який розпізнає текст вихідного зображення і в автоматичному режимі будує цифрову мапу. Результат роботи завантажується у вікно редагування, в якому користувач має можливість у ручному режимі відкоригувати текстову мапу. Редактор-маркер та редактор мап розроблені за допомогою середовища розробки Lazarus.