

**LIETUVOS ARCHYVŲ DEPARTAMENTAS PRIE LIETUVOS RESPUBLIKOS VYRIAUSYBĖS
ELEKTRONINIŲ DOKUMENTŲ SAUGOJIMO VALSTYBĖS ARCHYVUOSE INFORMACINĖS
SISTEMOS PROTOTIPO (BANDOMOJO PROJEKTO) SUKŪRIMAS**

**REKOMENDACIJOS
DĖL ELEKTRONINIO DOKUMENTO TURINIO ILGALAIKIO IR NUOLATINIO SAUGOJIMO
BEI PERŽIŪRAI INTERNETE TINKAMŲ (VAIZDO) FORMATŲ**

VILNIUS
2008

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Rekomendacijų rengėjai

Vardas Pavardė	Institucija
Adomas Birštunas	UAB „MitSoft“
Donatas Čiukšys	UAB „MitSoft“
Laimutis Ignatavičius	UAB „Sintagma“
Antanas Mitašiūnas	UAB „MitSoft“
Giedrius Noreikis	UAB „MitSoft“
Tadas Savičius	UAB „Sintagma“
Peep Uus	UAB „Sintagma“

Dokumento peržiūros istorija

Data	Versija	Aprašymas
2008-12-15	0.1	Lietuvos archyvų departamentui pateiktas pirminis variantas.
2008-12-19	0.2	Papildytas 3.4.1 skyrius.
2008-12-23	0.3	Atlikta dokumento korektūra, atsižvelgta į LAD pateiktas pastabas.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Turinys

1. Įžanga	4
1.1. Paskirtis	4
1.2. Santrumpos	4
2. Ilgalaikio saugojimo formatų palyginimas	5
2.3. PDF/A formato apžvalga	5
2.4. XPS formato apžvalga	7
2.5. TIFF formato apžvalga	8
2.6. Formatų palyginimas	8
2.7. Bandymų rezultatai	9
2.8. Interneto bendruomenės patirtis	10
2.9. Išvados	11
3. Peržiūrai internete tinkamų (vaizdo) formatų palyginimas	12
3.1. TIFF formato apžvalga	12
3.2. JPEG formato apžvalga	15
3.3. PNG formato apžvalga	18
3.4. Bandymų rezultatai	20
3.4.1. Sukurtų formato transformavimo priemonių išbandymas	20
3.4.2. Galimybių peržiūrėti formatus be papildomos programinės įrangos bandymai	21
3.4.3. Vaizdų saugojimo skirtingų parametrų JPEG ir PNG formatais bandymai	21
3.4.4. Vaizdų saugojimo skirtinga raiška bandymai	25
3.5. Formatų palyginimas	25
3.6. Interneto bendruomenės patirtis	27
3.7. Išvados	28
4. Rekomendacijos	30

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

1. Įžanga

1.1. Paskirtis

Šį dokumentą parengė Vykdytojas, įgyvendindamas 2008 m. rugpjūčio 18 d. sutartį Nr. VX-2008/08-LAD1/F8-13 (toliau – Sutartis), pasirašytą tarp Lietuvos archyvų departamento prie Lietuvos Respublikos Vyriausybės ir UAB „Sintagma“, veikiančios pagal Jungtinės veiklos sutartį, sudarytą 2008-06-12 tarp UAB „Sintagma“, UAB „MitSoft“ ir UAB „Etnomedijos intercentras“.

Rekomendacijose dėl elektroninio dokumento turinio ilgalaikio ir nuolatinio saugojimo bei peržiūrai internete tinkamų (vaizdo) formatų apibendrinami bandomojo projekto rezultatai ir Vykdytojo sukaupta formatų analizės, elektroninių dokumentų turinio transformavimo priemonių kūrimo ir išbandymo patirtis.

1.2. Santrumpos

Santrumpa	Santrumpos apibūdinimas
JPEG formatas	Fotografinių vaizdų saugojimo formatas, aprašytas standarte ISO/IEC 10918-1:1994 „Information technology. Digital compression and coding of continuous-tone still images: Requirements and guidelines“.
LAD	Lietuvos archyvų departamentas prie Lietuvos Respublikos Vyriausybės
ODF	Atvirieji biuro dokumento formatai, aprašyti standarte LST ISO/IEC 26300:2007 „Informacijos technologija. Atvirasis biuro dokumentų formatas v 1.0 (tapatus ISO/IEC 26300:2006)“.
OOXML	Atvirieji <i>Microsoft Office 2007</i> rinkmenų formatai, aprašyti standarte ISO/IEC DIS 29500 „Information technology. Office Open XML file formats“.
PDF/A formatas	Ilgalaikiam dokumentų saugojimui skirtas PDF formatas, aprašytas standarte ISO 19005-1:2005 „Document management. Electronic document file format for long-term preservation. Part 1: Use of PDF 1.4 (PDF/A-1)“.
PNG formatas	Perkeliamosios tinklo grafikos formatas, aprašytas standarte ISO/IEC 15948:2004 „Information technology. Computer graphics and image processing. Portable Network Graphics (PNG): Functional specification“.
TIFF formatas	Žymėtosios atvaizdų rinkmenos formatas, aprašytas standarte ISO 12639:2004. Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology (TIFF/IT) bei Adobe Developers Association specifikacijoje „TIFF Revision 6.0“.
Užsakovas	Lietuvos archyvų departamentas prie Lietuvos Respublikos Vyriausybės
Vykdytojas	UAB „Sintagma“, UAB „Mitsoft“ ir UAB „Etnomedijos intercentras“, veikiantys pagal Jungtinės veiklos sutartį, sudarytą 2008-06-12.
XPS formatas	„Microsoft“ korporacijos specifikaciją „XML Paper Specification“ atitinkantis formatas, žr. http://www.microsoft.com/whdc/xps/ .

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

2. Ilgalaikio saugojimo formatų palyginimas

Bandomojo projekto metu elektroninio dokumento turinio ilgalaikio ir nuolatinio saugojimo formatui parinkti buvo vertinami šie formatai:

PDF/A formatas	Ilgalaikiam dokumentų saugojimui skirtas PDF formatas, aprašytas standarte ISO 19005-1:2005 „Document management. Electronic document file format for long-term preservation. Part 1: Use of PDF 1.4 (PDF/A-1)“.
XPS formatas	„Microsoft“ korporacijos specifikaciją „XML Paper Specification“ atitinkantis formatas, žr. http://www.microsoft.com/whdc/xps/ .

Pastaba. Atliekant elektroninių dokumentų turinio transformavimo į TIFF formatą bandymus, buvo pastebėta, kad TIFF formatas dėl didelės rinkmenos apimties ir kitų priežasčių nėra tinkamas kaip vaizdo peržiūros internete formatas, tačiau jį naudojant išsaugomas patikimas elektroninio dokumento turinio atvaizdas. Dėl šios priežasties buvo analizuojama, ar TIFF formatas gali būti tinkamas elektroninio dokumento turinio ilgalaikiam ir nuolatiniam saugojimui.

2.3. PDF/A formato apžvalga

ISO 19005-1:2005 standarte aprašytas PDF/A formatas, skirtas elektroninių dokumentų ilgalaikiam saugojimui. Šio formato pagrindas yra PDF 1.4 specifikacija. PDF dokumentuose tekstinei informacijai atvaizduoti naudojami šriftai, kurie nebūtinai yra įtraukiami į šio formato rinkmeną, dokumento turinys gali būti šifruojamas, apsaugomas slaptažodžiais, o tai kelia rinkmenos perkėlimo problemų. Dokumento turinys gali būti priklausomas ir nuo išorinių nuorodų, kitų susijusių dokumentų, kurių dokumentą peržiūrint po kelerių metų gali būti nebeįmanoma surasti ir tai apribotų dokumento naudojimą.

Taikant minėtą standartą siekiama išspręsti anksčiau minėtas elektroninių dokumentų ilgalaikio saugojimo problemas, aprašant elektroninio dokumento formatui keliamus reikalavimus, kurie leidžia: 1) dokumentą papildyti metaduomenimis; 2) apibrėžti elektroninių dokumentų loginę struktūrą ir savybes. To rezultatas – PDF 1.4 formato pagrindu sukurtas naujas rinkmenos formatas PDF/A-1, kuris rinkmenoje talpina visą reikiamą informaciją, aprašo savo struktūrą ir yra labiau nepriklausomi nuo techninės įrangos negu įprastas PDF 1.4 formatas. Taip formatas tampa tinkamesnis ilgalaikiam dokumentų saugojimui.

Pagrindinis PDF/A skirtumas nuo PDF yra toks, kad PDF/A formato rinkmenoms keliami papildomi reikalavimai. Į PDF/A-1 formato rinkmeną privalo būti įtraukta:

- teksto rodymui reikalingi šriftai;
- nuo konkrečios įrangos nepriklausančios spalvos;
- XMP metaduomenys.

PDF/A-1 formato rinkmenos:

- neturi būti šifruojamos;
- neturi būti glaudinamos LZW algoritmu;
- neturi būti pildomos pridedamais dokumentais;
- neturi būti pildomos išorinėmis turinio nuorodomis;
- neturi turėti „PDF skaidrumo“ (angl. *PDF Transparency*) savybės;

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

- neturi turėti daugialypės terpės (angl. *multi-media*) elementų;
- neturi turėti *JavaScript* kodo.

Pagrindinės PDF/A-1 formato savybės:

- Nepriklausomas nuo techninės įrangos – naudojamas toks statinio vaizdo aprašas, kuris nėra priklausomas nuo techninės ir programinės įrangos, taip pat draudžiama formatą papildyti elementais, kurie nėra aprašyti PDF 1.4 specifikacijoje.
- Visą reikiamą informaciją turintis formatas – visa informacija ir duomenys, reikalingi PDF/A-1 formato turiniui atvaizduoti ir išspausdinti, turi būti saugoma formato rinkmenoje.
- Aprašantis dokumentą ir jo struktūrą – PDF/A-1 formato rinkmenoje privalo būti XMP metaduomenys, aprašantys į rinkmeną įraukiamus šriftus ir pagrindinius duomenis apie dokumentą, tokius kaip bylos identifikatorius, šriftų metaduomenys ir kt.
- Neapriboto naudojimo – neleidžiamas šifravimas ar apribojimas slaptažodžiais; jeigu to reikia, turi būti naudojamos kitos, išorinės priemonės..
- Viešo naudojimo – PDF/A-1 formato pagrindas yra viešai prieinama specifikacija ir daugybė programinių priemonių.
- Priimtinas – PDF/A-1 remiasi tokiomis technologijomis, kurios yra plačiai naudojamos visame pasaulyje; tik tokiu būdu galima tikėtis, kad formatą palaikančios programinės priemonės bus kuriamos tiek laiko, kiek jų reikės naudojimui.

Tačiau dokumento tinkamumas ilgalaikiam saugojimui priklauso ne vien tik nuo dokumento formato. Pavyzdžiui, dauguma šriftų, naudojamų tekstinei informacijai atvaizduoti, yra licencijuojami ir turi naudojimo, platinimo ir kitus apribojimus. PDF/A standartas reikalauja, kad šriftai būtų įtraukti į dokumento rinkmeną, todėl organizacijos, naudojančios PDF/A-1 formatus, turi imtis atsargumo priemonių, kad įsitikintų, ar turimos licencijos leidžia organizacijoje naudojamus šriftus įtraukti į tokias rinkmenas.

ISO standartas išskiria du suderinamumo su PDF/A-1 formatu lygius: A ir B lygiai, kurie iš esmės skiriasi teksto atvaizdavimui keliamais reikalavimais.

Siekiant PDF/A-1a suderinamumo (visiško suderinamumo) su standartu, reikalaujama, kad tekstas būtų pateikiamas tokia tvarka, kuria jis turėtų būti perskaitomas. Ši savybė (angl. *tagged PDF*) leidžia perskaityti tekstą riboto dydžio ekrane (pvz., mobiliajame telefone). Savybė gali būti naudinga dokumento teksto indeksavimui.

PDF/A-1b (minimalus) suderinamumo su standartu lygis reikalauja užtikrinti vizualų turinio (įskaitant tekstą) atvaizdavimą (pvz., kompiuterio ekrane), nepriklausomai nuo teksto pateikimo tvarkos. Tai negarantuoja, kad tekstas bus pateikiamas nuosekliai, tačiau bet kuriuo atveju jis bus teisingai pavaizduotas ekrane ar išspausdintas.

PDF 1.4 formatas, kurio pagrindu parengtas PDF/A-1 formatas, yra toks rinkmenos formatas, kuriame dokumentas atvaizduojamas nepriklausomai nuo kuriant, atvaizduojant dokumentą arba jį spausdinant naudotos programinės, techninės įrangos ir operacinės sistemos. PDF formato dokumentas sudarytas iš puslapių informaciją aprašančių objektų. Puslapiai ir kiti rodomieji elementai gali būti sudaryti iš teksto, grafikos ir paveikslėlių. PDF naudoja *PostScript* kalbą, kuri leidžia tekstą ir grafiką

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

aprašyti nepriklausančiu nuo įrangos ir raiškos būdu (skirtingai nuo taškinės grafikos vaizdų aprašymo formatų). PDF rinkmena yra sudaryta iš baitų (binarinė). Vaizdo ir kitų objektų informacija formate saugoma srovėmis (angl. *streams*), kurios informacija gali būti glaudinama informaciją neprarandančiu *deflate* algoritmu (naudojamu ZIP formate). Spalvoto arba nespalvoto vaizdo paveikslėliams glaudinti gali būti naudojami ir JPEG glaudinimo algoritmai, o vienspalviams paveikslėliams glaudinti – ICCT (3 arba 4), JBIG2 ir kiti algoritmai.

PDF formatas pradėtas kurti 1990 metais, pirmoji jo specifikacija paskelbta 1993 metais. PDF 1.4 formatas parengtas 2001 metais. Tuo metu pasaulyje buvo išplatinta daugiau kaip 200 mln. nemokamų programinės įrangos *Acrobat Reader* kopijų.

PDF/A-1 standartas patvirtintas 2005 metais. Šiuo metu rengiamas naujas, PDF/A-2 standartas, kuriame bus atsižvelgta į pagrindines naujoves, padarytas PDF 1.5, 1.6 ir 1.7 formatuose.

2.4. XPS formato apžvalga

XPS formatas yra 2006 metais „Microsoft“ korporacijos XML formato pagrindu sukurtas puslapio aprašymo formatas, kuriuo puslapio vaizdas aprašomas nepriklausomai nuo įrangos, reikalingos vaizdui parodyti ekrane ar jį išspausdinti ir nepriklausomai nuo vaizdo raiškos (skirtingai nuo taškinės grafikos vaizdų aprašymo formatų).

XPS formato rinkmena yra ZIP formato rinkmena, atitinkanti *Office Open XML Open Packaging Conventions* reikalavimus. Rinkmeną sudaro dalys (angl. *parts*) ir sąryšiai (angl. *relationships*). XPS rinkmena talpina dokumentų sekų, dokumentų, puslapių, šriftų, JPEG, PNG, TIFF arba *Windows Media Photo* formato paveikslėlių, nutolusių resursų, dokumento struktūros ir kitas dalis.

XPS dokumentui aprašyti naudojama XAML pagrindo kalba, naudojanti vektorinę grafiką dokumento elementams pavaizduoti. XPS rinkmeną sudaro kiekvieno dokumento puslapio, jame esančio teksto, naudojamų šriftų, taškinės grafikos vaizdų, dvimatės vektorinės grafikos ir teisių administravimo aprašai. XPS dokumento turinį galima tyrinėti, rinkmenos plėtinį tiesiog pervadinus plėtiniumi ZIP.

Kaip ir *Adobe PDF* formatas, XPS formatas užtikrina nesikeičiantį dokumento vaizdo išsaugojimą bet kuriame įrenginyje. Jei PDF formatas yra tiesiog *PostScript* kalbos objektų rinkinys, tai XPS formatas naujausiose *Windows* operacinėse sistemose (pvz., *Microsoft Vista*) yra skirtas naudoti kaip *PostScript* kalbą visiškai pakeičiantis formatas.

XPS, kaip ir PDF/A-1 formatas, reikalauja, kad į XPS rinkmeną būtų įtraukti visi teksto atvaizdavimui reikalingi šriftai, sudaromi spalvų aprašai ir metaduomenys.

Formatas neseniai sukurtas, todėl jis dar nėra labai paplitęs. „Microsoft“ korporacija yra sukūrusi priemones, leidžiančias naudoti formatą *Windows XP*, *Windows Server 2003* ir *Windows Vista* operacinėse sistemose. *Windows Vista* operacinėje sistemoje, ir, teigiama, kad kitose naujose *Windows* sistemose, ši formatą realizuojančios priemonės kompiuteryje bus įdiegiamos standartiškai, diegiant operacinę sistemą. *Windows Vista* operacinėje sistemoje naudojant *Internet Explorer 7* naršyklę XPS formatą galima peržiūrėti be papildomai įdiegtų programinių priemonių.

„Microsoft“ korporacija formatą pateikė standartizuoti.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

2.5. TIFF formato apžvalga

TIFF formato apžvalga pateikta 3.1 skyriuje.

2.6. Formatų palyginimas

Lentelėje pateiktas PDF/A-1 ir XPS formatų palyginimas:

Kriterijus	PDF/A-1	XPS
XML pagrindo formatas	Ne	Taip
Vektorinė grafika	Taip, <i>PostScript</i> kalba	Taip, visiškai pakeičianti <i>PostScript</i> kalbą
Perkeliamas formatas	Taip	Taip
Informacijos glaudinimas	Taip	Taip
Greitas puslapio po puslapio peržiūrėjimas siunčiantis internetu	Taip	Taip
Užtikrina dokumento atvaizdo išsaugojimą spausdinant	Taip	Taip
Naudoja CMYK spalvų paletę profesionaliam spausdinimui	Taip	Taip
Palaiko JPEG paveikslėlius	Taip	Taip
Palaiko PNG paveikslėlius	Ne	Taip
Palaiko TIFF paveikslėlius	Ne	Taip
RGB spalvų paletė	Taip	Taip
CMYK spalvų paletė	Taip	Taip
Nespalvotas vaizdas	Taip	Taip
Palaikomi TrueType šriftai	Taip	Taip
Palaikomi Type1 (PostScript) šriftai	Taip	Ne
Palaikomi vektoriniai šriftai	Taip	Ne
Nacionalinių simbolių palaikymas	Taip (tik <i>Unicode</i>)	Taip
Visiškas šrifto įtraukimas į rinkmeną	Taip	Taip
Šrifto dalies įtraukimas į rinkmeną	Taip	Taip
Apsauga nuo šriftų kopijavimo už rinkmenos	Taip	Taip (iš dalies)
Rekomenduotinas nuotoliniam spausdinimui	Ne	Taip
Rekomenduotinas dokumentų archyviniam saugojimui	Taip	Taip
Rekomenduotinas dokumentų atsisiuntimui internetu	Taip	Ne (nepaplites)

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

2.7. Bandymų rezultatai

Sukurtų transformavimo priemonių bandymų metu pastebėta, kad transformuojant 40 ADOC specifikaciją atitinkančių dokumentų turinius (bendras dokumentų dydis – 23 MB), turinių transformavimas į XPS formatą truko beveik 7 minutes, rezultatas – daugiau kaip 98 MB. Transformuojant tą pačią imtį dokumentų į PDF/A-1 formatą, transformavimas truko apie 5 minutes, rezultatas – 50 MB.

Atliekant pakartotinius bandymus su didesniu dokumentų skaičiumi, 1000 dokumentų (bendra apimtis – 545 MB, turinio rinkmenų skaičius – 1928), buvo gauti šie turinio transformavimo rezultatai:

Palyginamasis kriterijus	XPS	PDF/A-1
Transformavimo trukmė (min.)	169	90,2
Bendras rezultatų dydis (MB)	2507,8	916

Elektroninių dokumentų transformavimas į PDF/A formatą trunka trumpiau, transformavimo rezultatų apimtis yra 2,7 kartų mažesnė, palyginti su el. dokumentų turinio transformavimu į XPS formatą rezultatais.

Vertinant atskirų turinio rinkmenų transformavimo greitį, pastebima, kad ODF turinio rinkmenų (*OpenDocument* rašyklės, skaičiuoklės ir pateikčių programomis sukurtų rinkmenų) transformavimas į XPS formatą trunka ilgiau nei rinkmenų transformavimas į PDF/A formatą. Taip yra todėl, kad XPS formatas yra naujas ir retai naudojamas kitose programinėse priemonėse, išskyrus tas, kurios yra sukurtos „Microsoft“ korporacijos. ODF formatų redaktoriai šiuo metu teikia galimybę išsaugoti turinio rinkmenas tik PDF/A formatu. Transformavimui į XPS formatą naudojama turinio spausdinimo *Microsoft XPS Document Writer* programine įranga technologija nepasižymi didele XPS rinkmenų sudarymo sparta (spausdinimo sparta nepasižymi ir pati spausdinimo technologija, kadangi spausdinimo įrenginiai laikomi didele sparta nepasižyminčiais įrenginiais).

Buvo pastebėta, kad transformuojant *Microsoft PowerPoint* pateikčių *pptx* ar *ppsx* formato rinkmenas į XPS formato turinį tiek programinės įrangos teikiama išsaugojimo XPS formatu funkcijomis, tiek pateiktį spausdinant *Microsoft XPS Document Writer* spausdintuvu, dokumento turinys iškraipomas. Keliose vietose pastebėta teksto iškraipymų (pasislinkimų), kitose skaidrėse matyti teksto šešėliai, nors peržiūrint originalų dokumentą pateikties peržiūros priemonėmis jų nėra matyti. Darytina išvada, kad problemų kyla dėl transformavimo į XPS formatą metu programinės priemonės nekorektiškai apdorojamų šriftų. Šių problemų nepastebėta pateikčių turinį išsaugojus PDF/A formatu. Turinio iškraipymai taip pat buvo pastebėti transformuojant *OpenOffice Calc* (*ods*) rinkmenas į XPS formatą. Pastarųjų nepastebėta rinkmenas transformavus į PDF/A formatą.

Atlikus skirtingų formatų turinio naudojamos bandymus, pastebėta, kad *Linux Fedora 9* operacinėje sistemoje naudojant programinę įrangą *Okular* XPS formato turinio rinkmenos rodomos iškraipytos. Kitų nemokamų priemonių XPS formato turinui peržiūrėti šioje operacinėje sistemoje nėra. Rinkoje galima įsigyti keletą komercinių priemonių, kuriomis galima peržiūrėti XPS turinį *Linux* operacinėje sistemoje. Problemų su PDF/A formato dokumentų turinio peržiūra *Windows XP*, *Windows 2003 Server*, *Windows Vista Business* ir *Linux Fedora 9* operacinėse sistemose nebuvo pastebėta.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

2.8. Interneto bendruomenės patirtis

Šaltinis: <http://www.amyuni.com/downloads/AmyuniWhitepaper-PDFvsXPS.pdf>. „PDF ar XPS: Pasirinkite geriausią dokumento formatą jūsų programoms“, Dany Amiouny, Amyuni Technologies.

„Pastebėta, kad XPS formatas yra rimtas PDF formato konkurentas ir jo privalumas yra tas, kad kuriant formatą buvo daug ko pasimokyta iš PDF formato, kuris ilgą laiką buvo kurtas ir keliomis iteracijomis tobulintas. Tačiau PDF formatas vis dar išlaiko pasirinkimo pirmumo teisę. Kol rinkoje yra milijonai PDF formato dokumentų ir šimtai PDF formato priemonių, PDF formatas dar ilgai bus palaikomas.“

Šaltinis: <http://labnol.blogspot.com/2006/01/adobe-pdf-vs-microsoft-xps-xml-paper.html>. „Adobe PDF ir Microsoft XPS (XML Paper Specification)“, Amit Agarwal, 2006.

„Microsoft XPS formato sėkmė labiausiai priklausys nuo to, kaip greitai programinės įrangos ir spausdintuvų gamintojai nutars kurti XPS formatą naudojančias tvarkykles ir įrenginius. <.> XPS formatas rimtai pretenduoja užimti PDF formato rinką, tačiau Microsoft Windows Vista yra išleista vos tik prieš keletą mėnesių, o „Adobe“ korporacija nėra linkusi sėdėti rankas sudėjusi.“

Šaltinis: <http://www.informationzen.org/group/pdf/forum/topics/2043787:Topic:1383>. Diskusija: „Kokie PDF, XPS ir OOXML formatų skirtumai?“, Scott Mackey komentaras, 2008-10-06.

„PDF ir XPS formatai leidžia platinti ir keistis dokumentais, kurių vaizdas nesiskiria nuo originalaus dokumento. Tačiau PDF formatas yra labiau subrendęs ir funkcionalus. PDF bendruomenė siūlo platų ir vis augantį sukurtų priemonių asortimentą. <.> XPS formatas yra naujas, ir kaip visada, jam teks nueiti daugumos „naujų“ formatų pripažinimo kelią (nepaisant to, kad turi solidų rėmėją ☺).“

Šaltinis: <http://www.informationzen.org/group/pdf/forum/topics/2043787:Topic:1383>. Diskusija: „Kokie PDF, XPS ir OOXML formatų skirtumai?“, Dwight Kelly komentaras, 2008-09-26.

„PDF yra specialus „Adobe“ korporacijos sukurtas binarinės rinkmenos formatas. XPS naudoja XML ir ZIP formatus. Tiesa, „Adobe“ korporacija yra sukūrusi XML pagrindo PDF formatą, vadinamą MARS. PDF formatas metų metus buvo tobulinamas, kad būtų suderinamas su beveik visomis dokumentų rūšimis. XPS formatą galima naudoti tik keičiantis ribotos rūšies el. dokumentais ir el. parašais.“

Pagrindinis XPS formato privalumas prieš PDF yra kaina, kadangi XPS priemonės *Windows XP* ir *Vista* operacinėse sistemose nekainuoja (įdiegiamos standartiškai arba naujinant sistemą). Tačiau esamos *Microsoft XPS Viewer* priemonės neturi svarbių funkcijų ir veikia labai lėtai.“

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

2.9. Išvados

XPS ir PDF/A-1 formatų palyginamoji analizė parodė, kad abu formatai (nekreipiant dėmesio į formatus realizuojančias priemones) yra tinkamai elektroninių dokumentų turinių atvaizdo ilgalaikiam saugojimui, kadangi jie yra apimantys visa ko reikia norint atvaizduoti informaciją ar ją išspausdinti, abu šie formatai informaciją saugo taikydami panašius principus ir yra laikomi tarpusavyje konkuruojančiais.

Nors abu formatai yra neseniai sukurti, tačiau jų pagrindas skiriasi. PDF/A-1 formato pagrindas yra interneto bendruomenės pripažintas ir plačiai naudojamas PDF 1.4 versijos formatas. Standartas tik apibrėžia reikalavimus, keliamus formato rinkmenos sudarymui, siekiant užtikrinti ilgalaikį dokumentų saugojimą, todėl priemonės, kuriomis galima peržiūrėti ar kitaip naudoti PDF 1.4 formato turinį, iš esmės yra tinkamos peržiūrėti ir PDF/A-1 formato turinį.

Tuo tarpu XPS formatas yra visiškai naujas, jį palaikančių priemonių yra nedaug, o pastarosios veikia dar ne taip gerai, kaip norėtųsi, nes bandymų metu buvo pastebėta, kad į šį formatą transformuotas vaizdas kartais iškraipomas. Ar šis formatas bus plačiai pripažintas, galima bus įsitikinti po 3-5 metų, kai paaiškės, ar standartas bus sertifikuotas ISO, ir kiek bus sukurta formatą realizuojančių programinių priemonių. Lyginant PDF/A-1 formatą su XPS formatu, šiuo metu elektroninių dokumentų ilgalaikiam saugojimui rekomenduotina rinktis PDF/A-1 formatą.

PDF/A-1 standartas apibrėžia du suderinamumo su standartu lygmenis: visiškas suderinamumas (PDF/A-1a) ir minimalus suderinamumas (PDF/A-1b). Abu šie lygmenys užtikrina vizualų turinio atvaizdą, tačiau jeigu turinio transformavimo programinė priemonė leidžia pasirinkti suderinamumo su standartu lygmenį, rekomenduotina rinktis PDF/A-1a.

Atlikus elektroninio dokumento turinio atvaizdo išsaugojimo skirtingos raiškos taškinės grafikos vaizdu bandymus (žr. 3.4.4 skyrių), nustatyta, kad kokybiškiausiai ekrane yra perskaitomos nuo vaizdo raiškos nepriklausančios PDF/A-1 formato rinkmenos. TIFF formato rinkmenos yra neproporcingai didelės (nuo 4 iki 12 kartų didesnės), palyginti su PDF/A-1 formato rinkmenomis.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

3. Peržiūrai internete tinkamų (vaizdo) formatų palyginimas

Bandomojo projekto metu buvo vertinami šie elektroninio dokumento turinio atvaizdavimo internete (vaizdo) formatai:

TIFF formatas	Žymėtosios atvaizdų rinkmenos formatas, aprašytas standarte ISO 12639:2004. „Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology (TIFF/IT)“ ir Adobe Developers Association specifikacijoje „TIFF Revision 6.0“.
JPEG formatas	Fotografinių vaizdų saugojimo formatas, aprašytas standarte ISO/IEC 10918-1:1994 „Information technology. Digital compression and coding of continuous-tone still images: Requirements and guidelines“.
PNG formatas	Perkeliamosios tinklo grafikos formatas, aprašytas standarte ISO/IEC 15948:2004 „Information technology. Computer graphics and image processing. Portable Network Graphics (PNG): Functional specification“.

Kaip koncepcinę alternatyvą taip pat tyrėme, ar gali būti ilgalaikiam saugojimui parinkti formatai PDF/A-1 ir XPS nagrinėjami kaip peržiūrai internete tinkami formatai.

3.1. TIFF formato apžvalga

Žymėtosios atvaizdų rinkmenos (TIFF) formatas yra taškinės grafikos vaizdų saugojimo formatas. Jis leidžia išsaugoti didelius ir aukštos kokybės vaizdus, neprarandant vaizdų kokybės. Formatas vienoje rinkmenoje leidžia saugoti keletą paveikslėlių, tačiau rinkmenos dydis neturi viršyti 4 GB.

Saugomos vaizdo informacijos dydis priklauso nuo paveikslėlio dydžio, raiškos, naudojamų spalvų ir pasirinkto vaizdinės informacijos glaudinimo algoritmo.

Priklausomai nuo pasirinktos raiškos (vaizdinės informacijos taškų skaičius viename colyje), vaizdas skeneryje yra skaidomas į pikselių (spalvinių ar nespalvinių taškų) masyvą (matricą), kuri sudaro eilutės ir stulpeliai (žr. 1 paveikslą).

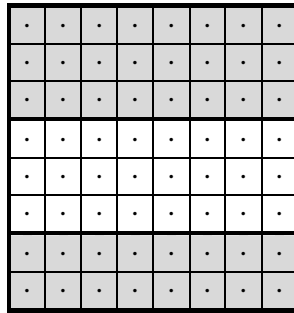
.
.
.
.
.
.
.
.

1 paveikslas. Taškinės grafikos vaizdo matrica

TIFF formate, naudojant specialias žymas, nurodomas taškinės grafikos vaizdo eilučių ir stulpelių skaičius ir vaizdo raiška. Saugomo vaizdo atvaizdavimui spartinti didesnis paveikslėlis gali

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

būti skaidomas į horizontalias juostas (angl. *strips*), apimančias nurodytą taškinės grafikos vaizdo matricos eilučių skaičių. Kiekviena juosta (žr. 2 paveikslą) TIFF formate yra saugoma atskirai ir gali būti naudojama rodyti vaizdą dalimis.



2 paveikslas. 3-jų eilučių pločio taškinės grafikos vaizdo juostos

Taškinės grafikos vaizdo juostoje vaizdinė informacija saugoma, iš eilės surašant pikselių masyvą.

TIFF formatas leidžia saugoti vienspalvius, nespalvotus ir spalvotus vaizdus.

Vienas vienspalvio vaizdo pikselis įgyja reikšmę 0 arba 1, kurios, priklausomai nuo fotometrinio interpretavimo (angl. *Photometric Interpretation*) žymos reikšmės, reiškia juodą arba baltą spalvą. Kitaip tariant, vienam vienspalviui pikseliui saugoti pakanka vieno bito, 8 pikseliams saugoti – vieno baito.

Nespalvotam vaizdui saugoti, priklausomai nuo atspalvių skaičiaus, reikia 4 bitų (16 atspalvių) arba 8 bitų (256 atspalviai) kiekvieno pikselio informacijai aprašyti.

Spalvoto vaizdo saugojimas priklauso nuo naudojamos vaizdo spalvų paletės. Jei naudojama RGB spalvų paletė, naudojami 3 baitai vienam pikseliui aprašyti, jei CMYK spalvų paletė – 4 baitai kiekvienam pikseliui aprašyti ir t.t.

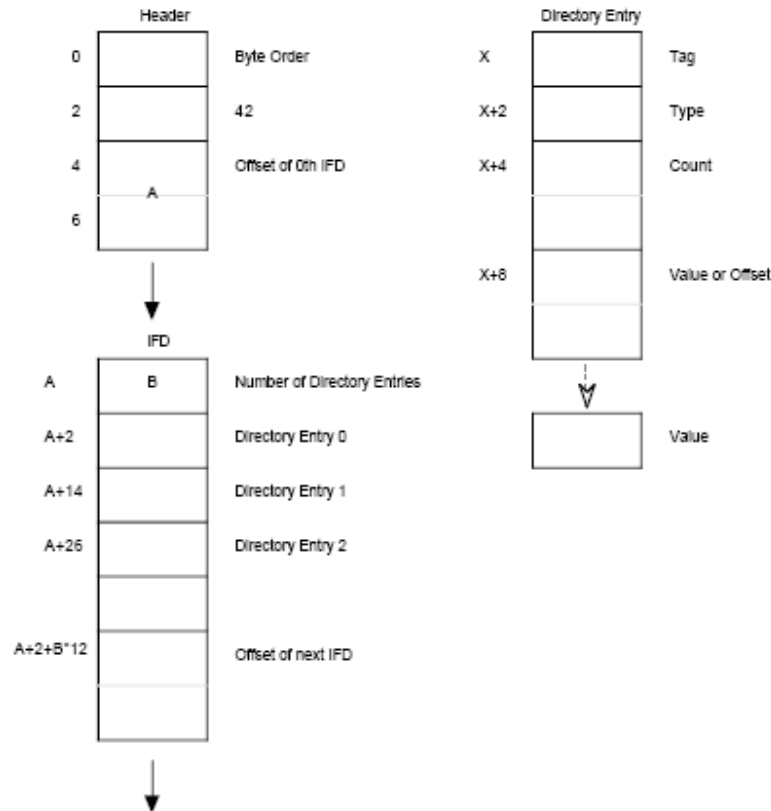
Vaizdinė informacija TIFF formate gali būti saugoma neglaudinta arba glaudinta, taikant glaudinimo algoritmus: CCITT Huffmano kodo arba LZW glaudinimo algoritmus (nuo 1992 metų naudojamas ir JPEG formate naudojamas glaudinimo algoritmas).

Vaizdinė informacija ir TIFF žymomis aprašyta vaizdo informacija talpinama į vadinamąjį paveikslėlių rinkmenos katalogą (IFD). TIFF rinkmeną gali sudaryti keli paveikslėlių rinkmenos katalogai, t.y. keli paveikslėliai (žr. 3 paveikslą). Vaizdą aprašanti informacija saugoma katalogo įrašuose (angl. *Directory entry*), kurį sudaro žyma (angl. *tag*), žymos reikšmės tipas (angl. *type*) – baitas, skaičius, tekstas – ir reikšmė (angl. *value*).

TIFF specifikacijoje apibrėžtos pagrindinės vaizdinei informacijai aprašyti reikalingos žymos, tačiau nedraudžiama plėsti formatą ir naudoti papildomas žymas. Reikalaujama, kad TIFF formatą skaitanti programinė įranga neatpažintąs TIFF žymas ignoruotų.

Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų



3 paveikslas. TIFF formato rinkmenos struktūra.

TIFF formatą 1986 metais sukūrė „Aldus“ korporacija, kuri kartu su „Microsoft“ korporacija pasiūlė skenerių ir spausdintuvų gamintojams susitarti dėl bendro ir universalios skenuoto vaizdo saugojimo formato, kad kiekvienas skenerių ar spausdintuvų gamintojas nekurtų savo saugojimo formato. Pirmoji TIFF formato specifikacija buvo paskelbta tik po gausybės susitikimų su skenerių ir programinės įrangos gamintojais ir jai buvo suteiktas 3.0 versijos numeris. 1987 metais su nedideliais patobulinimais buvo išleista 4.0 specifikacijos versija, kuri 1988 metų spalio mėnesį papildyta galimybe saugoti spalvotus vaizdus ir LZW glaudinimo algoritmu (kuris 1995 metais „Unisys“ korporacijos buvo užpatentuotas; patentų teisės galiojo iki 2004 metų).

TIFF standarte yra numatyta galimybė plėsti formatą, naudojant naujas žymas. Kadangi standartas nereguliuoja šių žymų įrašymo algoritmo, vos ne kiekvienas techninės ar programinės įrangos gamintojas įrašydavo savo žymas ir jas naudodavo. Taip formatas tapo probleminiu, nes kilo sąveikumo problema: rinkoje atsirado tūkstančiai TIFF atmainų ir nebuvo nė vienos programinės priemonės, kuri visas tas atmainas palaikytų.

„Adobe“ korporacija, „Aldus“ korporacijos teisių perėmėja, 1992 metais paskelbė TIFF 6.0 specifikaciją, kurioje atsižvelgė į tuo metu rinkoje naudotas populiariausias TIFF formato atmainas, ir ėmė taikyti žymų registravimo mechanizmą. Specifikacija buvo papildyta galimybėmis aprašyti

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

vaizdus CMYK, YCbCr ir kitomis spalvų paletėmis, pradėtas naudoti JPEG formate naudojamas vaizdų glaudinimo algoritmas.

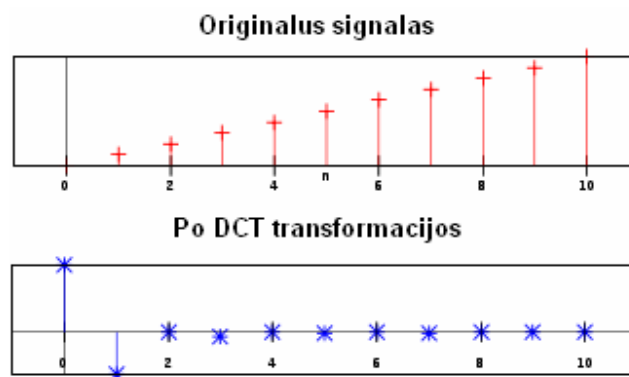
Siekiant TIFF formatą padaryti mažiau komplikuoju, TIFF 6.0 specifikacijos pagrindu buvo parengta keletas skirtingų standartų: ISO 12234-2:2001 *Electronic still-picture imaging -- Removable memory -- Part 2: TIFF/EP image data format*; ISO 12639:2004 *Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology* (TIFF/IT).

TIFF formatas yra vienas seniausių vaizdų formatų ir dauguma vaizdus apdorojančių programinių priemonių jį realizuoja. Tačiau kadangi TIFF formate nebuvo numatyta efektyvių vaizdų glaudinančių algoritmų (LZW algoritmas buvo užpatentuotas 1995 metais), vaizdinė informacija jame dažniausiai saugoma nesuspausta ir dėl to nėra tinkama keitimuisi internete. Dėl šios priežasties populiariausios interneto naršyklės *Internet Explorer*, *Mozilla Firefox* ir *Opera* šio formato pačios nerealizuoja – jos pasiūlo pasirinkti programinę priemonę rinkmenos turiniui peržiūrėti.

3.2. JPEG formato apžvalga

Fotografinių vaizdų saugojimo (JPEG) formatas yra taškinės grafikos vaizdų saugojimo formatas. Šis formatas labiausiai tinkamas saugoti fotografinius vaizdus. Formatas apibrėžia tokį informaciją prarandantį glaudinimo algoritmą, kad net iki 10% kokybės suglaudintas vaizdas yra pakankamai raiškus ir atpažįstamas. Tačiau formatas nėra labai tinkamas brėžiniams ar tekstinei informacijai saugoti, kadangi glaudinat vaizdą mažėja linijų, tekstinių simbolių ir kitų kontrastinių kontūrų raiška. Tai „užkoduota“ vaizdo glaudinimo algoritme.

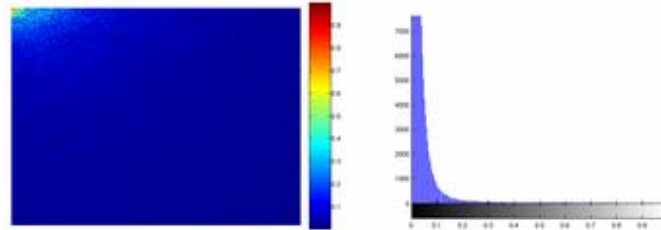
JPEG formato pagrindas – vaizdo glaudinimo algoritmas, pagrįstas diskretine kosinų transformacija (*DCT*), t.y. diskretine Furjė transformacija, kurioje naudojami tik realieji skaičiai. Furjė transformacija naudojama apdorojant analoginius signalus ir pasižymi signalų „energijos glaudinimo“ savybe, aprašančia signalų tendenciją: kuo signalas monotoniškesnis, tuo mažiau reikšmių reikia signalui aprašyti (žr. 4 paveikslą). Informacija, turinti daugiau pasikartojimų, glaudinama efektyviau. Atitinkamai efektyviau glaudinamas ir mažiau kontrastų turintis vaizdas (žr. 5 paveikslą).



4 paveikslas. DCT transformavimo rezultato pavyzdys

Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų

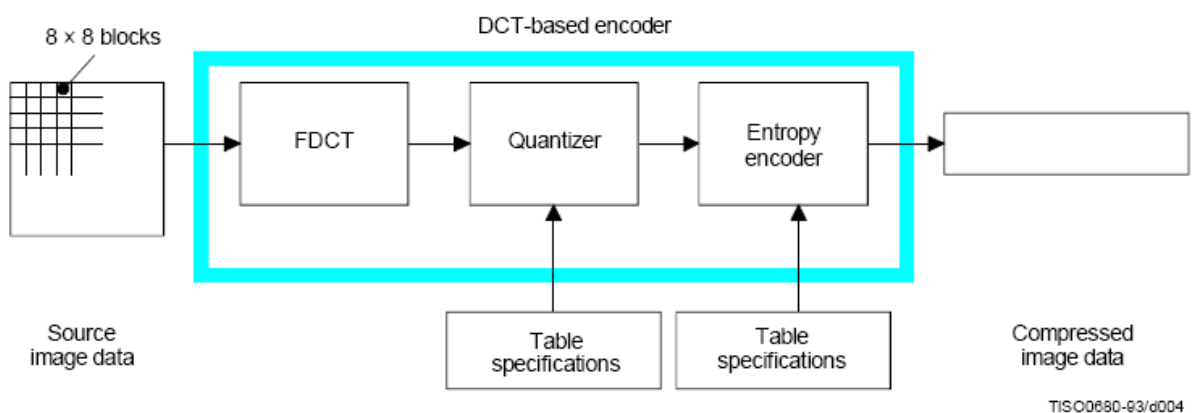


5 paveikslas. Vaizdinės informacijos DCT transformavimo rezultato pavyzdys.

Vaizdas transformuojamas 8x8 pikselių matricomis (žr. 1 paveikslą). Kiekvienos vaizdo matricos pikseliui atliekama (žr. 6 paveikslą) tiesioginė diskretinė kosinų transformacija (*FDCT*), paverčianti taškinės grafikos matricą 64 reikšmių aibe, vadinama DCT koeficientais. Kiekvienas DCT koeficientas yra kvantizuojamas (apribojamos koeficiento reikšmės, dalijant reikšmes iš parametru ir taikant informaciją prarandantį rezultato apvalinimą iki sveikojo skaičiaus) naudojant glaudinimo parametrus. Tuomet zig-zag sekos būdu (žr. 7 paveikslą) sudaroma 64 reikšmių entropijos (sistemos tvarkos matas, rodantis, kaip arti pusiausvyros yra termodinaminė sistema) seka. Entropijos reikšmių sekos gali būti glaudinamos pasirinktu algoritmu: Huffmano kodavimo algoritmu arba aritmetinio kodavimo algoritmu.

JPEG standarte apibrėžti 4 kodavimo operacijų režimai:

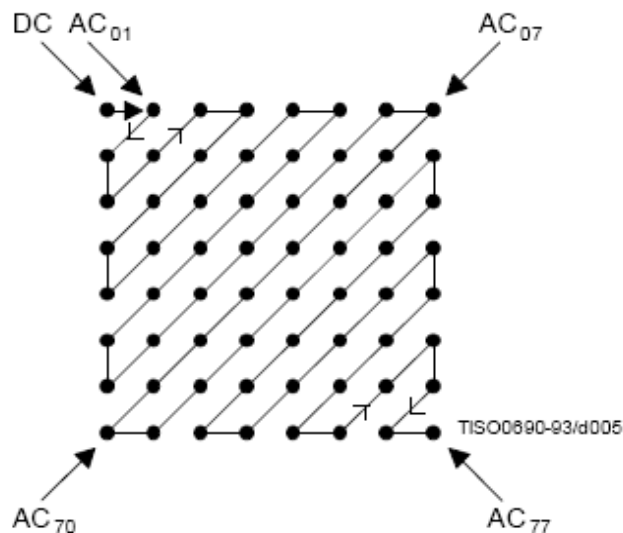
- nuoseklus DCT transformacijos režimas, kurio metu 8x8 pikselių matricos transformuojamos paėiliui iš kairės į dešinę ir iš viršaus žemyn;
- progresinis DCT transformacijos režimas, kurio metu atliekama keletas vaizdo skenavimų tam tikru algoritmu. Šis metodas sudaro galimybę anksčiau pamatyti ekrane bendrus vaizdo kontūrus, kurie vėliau yra palaipsniui papildomi detalėmis iki pilno vaizdo;
- vaizdinės informacijos neprarandantis režimas; ir
- hierarchinis režimas, kuomet sudaromi keli hierarchiniai to paties vaizdo kadrai, pritaikyti peržiūrėti vaizdą skirtinga raiška.



6 paveikslas. DCT transformavimo algoritmas

Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų



7 paveikslas. Zig-zag seka, kuria remiantis sudaroma entropijos reikšmių seka

Saugomai vaizdinei informacijai atvaizduoti naudojamas atvirkštinis DCT transformavimo algoritmas (*IDCT*).

Vaizdo glaudinimo metu kiekviena 8x8 pikselių matricos taško spalva pakeičiama entropijos reikšme, nustatyta atsižvelgiant į kitų matricoje esančių taškų spalvas. Kai vaizdas turi ryškių kontrastų (pavyzdžiui, juoda linija baltame fone), vaizdo išglaudinimo metu linijoje ir šalia jos esantiems taškams suteikiama spalva yra „tarpinė“ tarp juodos ir baltos spalvų, t.y. viena iš pilkos spalvos atspalvių. Tai paaiškina, kodėl algoritmas nėra tinkamas linijų, tekstinių simbolių ir kitų kontrastinių kontūrų turintiems vaizdams saugoti: linijos vizualiai „išblunka“, „išstęžta“, fonas, kuriame parašytas tekstas, „susipurvina“ (žr. 8 paveikslą). Tai yra „užkoduota“ vaizdo glaudinimo algoritme – nuo formato atvaizdavimo priemonių tai nepriklauso.



8 paveikslas. Tekstas JPEG formatu (kokybė 10%).

Tačiau JPEG puikiai tinkamas saugoti fotografiniams vaizdams, kurie yra mažiau kontrastingi – kiekvieno pasirinkto vaizdo taško spalva panaši į aplinkinių taškų spalvas. Šiuo metu JPEG formatas ir jo atmainos plačiai naudojamos vaizdams saugoti skaitmeniniuose fotoaparatuose.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

JPEG formatą sukūrė 1986 metais suburta Jungtinė fotografijos ekspertų grupė, kuri 1992 metais parengė standartą, 1994 metais patvirtintą ISO/IEC 10918-1:1994. Tuo metu jau buvo sukurtas TIFF formatas.

JPEG vaizdų saugojimo formatas, apibrėžtas standarto B priede (vadinamas JIF), yra retai naudojamas visų pirma dėl sunkumų realizuoti visus standarte numatytus atvejus. Siekiant išspręsti šias problemas, buvo sukurtas keletas naujų formatų: *JPEG File Interchange Format* (JFIF), *Exchangeable image file format* (Exif), kt.

Sumaišties JPEG istorijoje 2002 metais įnešė kompanija Forgent Networks, pareiškusi patento teises į JPEG formatą, ir šios teisės buvo pripažintos. Interneto bendruomenė, reaguodama į tai sukūrė keletą naujų JPEG formato atmainų, iš kurių žinomiausias yra JPEG2000 formatas. 2006 metais teisme buvo įrodyta, kad kompanija, valdanti JPEG formato patentą, žinojo apie JPEG standartizavimo procedūras, ir laiku, kaip priklausė, nepareiškė patento teisių į standartizuojamą formatą, todėl patento teisės kompanijai buvo panaikintos.

3.3. PNG formato apžvalga

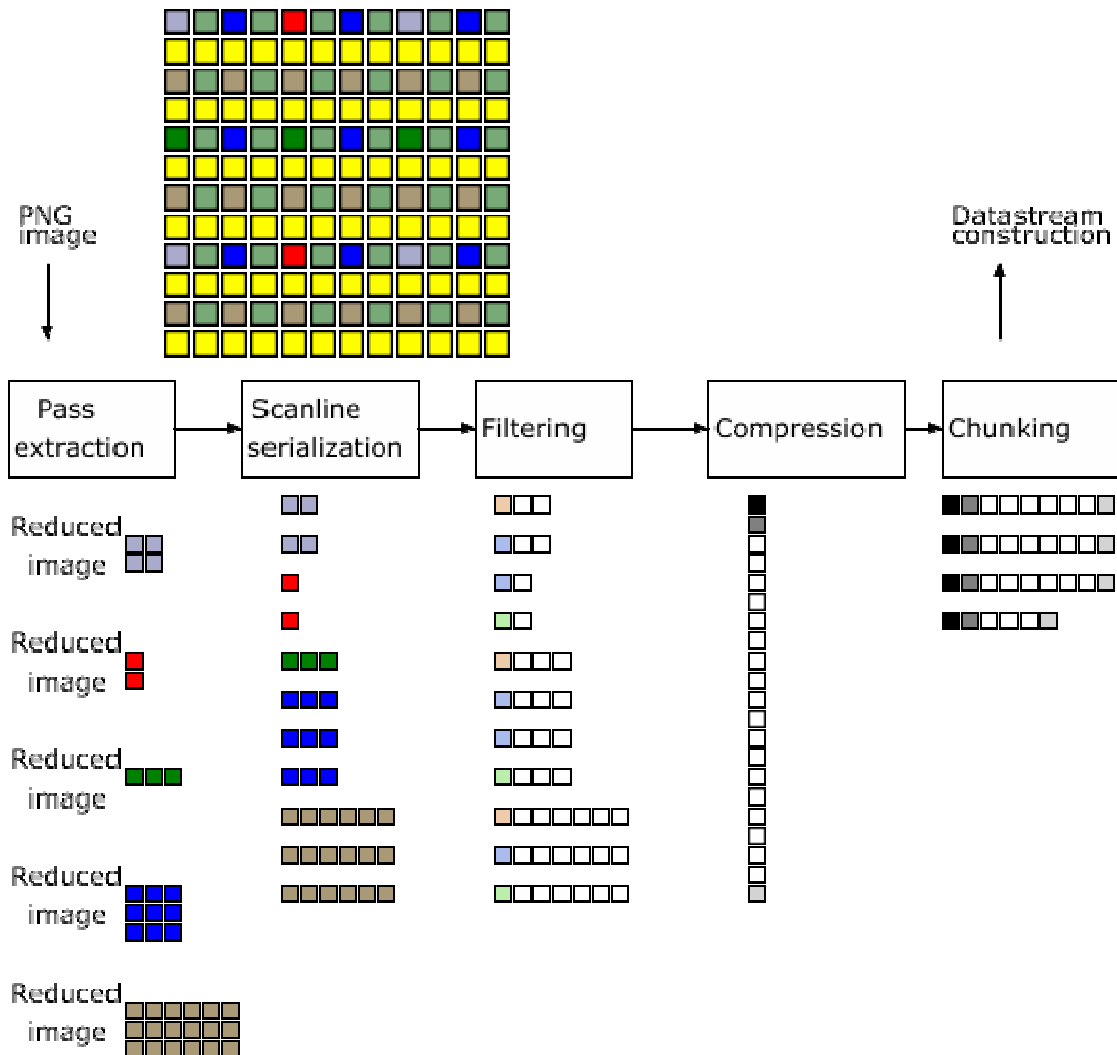
Perkeliamosios tinklo grafikos (PNG) formatas yra taškinės grafikos vaizdų saugojimo formatas. Šis formatas pasižymi plečiamumu, perkeliamumu tarp skirtingos techninės ir programinės įrangos, ir puikiu taškinės grafikos vaizdo glaudinimu. PNG formatas yra tobulesnis už GIF formatą ir nesaugomas patento teisių (skirtingai nuo GIF formato). Šis formatas tinkamas pakeisti TIFF formatą dauguma pastarojo formato panaudos atvejų.

Pagrindiniai deklaruojami formato sukūrimo tikslai – vaizdo perkeliamumas, pilnumas (galima saugoti įvairių spalvų vaizdus su permatomo fono galimybėmis, įskaitant ir tekstinius komentarus), serijinis vaizdinės informacijos kodavimas, progresinis vaizdų pateikimas, klaidų apdorojimas (galima nustatyti, kada formatas yra pažeistas), vaizdų apdorojimo greitis, glaudinimo efektyvumas, neprarandant vaizdinės informacijos, paprastumas (nesudėtinga sukurti formatą realizuojančias priemones) ir kt.

Taškinės grafikos vaizdas pirmiausia yra suskaidomas į spalvų kanalus: RGB paletės spalvų vaizdai skaidomi į tris kanalus (raudonos, žalios ir mėlynos spalvos), RGBA paletės spalvų vaizdai – į keturis, nespaltoto – į vieną arba du kanalus. Tuomet (žr. 9 paveikslą) taškinės grafikos vaizdas skaidomas (angl. *pass extraction*) į mažesnius progresinius vaizdus (PNG standarte aprašyti du skirtingi progresinių vaizdų sudarymo algoritmai). Gautos vaizdų pikselių eilutės paverčiamos baitų sekomis (angl. *scanline serialization*), kurių reikšmės specialiu algoritmu yra filtruojamos (angl. *filtering*), taip sekas parengiant glaudinimui (PNG standarte aprašytas 1 filtravimo metodas su 5 skirtingais parametrais (algoritmais), naudojamais apskaičiuojant reikšmių skirtumus, lyginant su gretimų taškų reikšmėmis). Gauta reikšmių seka yra glaudinama (angl. *compression*) (PNG standarte minimas informacijos neprarandantis *deflate* glaudinimo algoritmas, naudojamas ZIP formate). Suglaudinta vaizdo informacija skaidoma į sekas (angl. *chunking*), apskaičiuojama kiekvienos iš jų kontrolinė suma (CRC), ir sudaroma PNG rinkmena.

Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų



9 paveikslas. Taškinės grafikos vaizdo kodavimo PNG formatu seka.

PNG formatas pradėtas kurti 1995 metais, kai „Unisys“ korporacija įgijo teises į LZW glaudinimo algoritmo, naudojamo GIF formate, patentą. Tuo metu jau buvo sukurti TIFF ir JPEG standartai. GIF formatas taip pat turėjo 256 spalvų gylio apribojimą, kurio norėta išvengti kuriant naują formatą. 1996 metais buvo paskelbta pirmoji PNG formato versija, kuri 1996-10-01 tapo W3C rekomendacija (RFC 2083). 2003 metais buvo išleistas standartas ISO/IEC 15948:2003, 2004 metais atnaujintas (ISO/IEC 15948:2004).

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

3.4. Bandymų rezultatai

3.4.1. Sukurtų formato transformavimo priemonių išbandymas

Sukurtų transformavimo priemonių bandymų metu pastebėta, kad transformuojant 40 ADOC specifikaciją atitinkančių dokumentų turinius (bendras dokumentų dydis – 23 MB):

Palyginamasis kriterijus	TIFF	JPEG	PNG
Transformavimo trukmė (min.)	15,3	27,5	27,0
Bendras rezultatų dydis (MB)	986,69	1047,9	543,7

Palyginamojoje lentelėje matyti, kad mažiausio dydžio rinkmenos (beveik dvigubai mažesnės už JPEG) yra PNG formato rinkmenos. Transformavimo trukmė buvo mažiausia transformuojant į TIFF formatą. JPEG ir PNG formatai gaminami iš TIFF formato rinkmenų.

Atlikus pakartotinius bandymus su didesne imtimi (1000 el. dokumentų, bendras dokumentų dydis 328,7 MB, bendras turinio rinkmenų skaičius 1519; vaizdo raiška 300x300 dpi), gauti rezultatai:

Palyginamasis kriterijus	TIFF(LZW)	JPEG(80%)	PNG
Transformavimo trukmė (min.)	140	229	225
Bendras rezultatų dydis (MB)	9886,7	5897	4237,1
<i>Rezultato dydis (MB) 100 MB pradinių duomenų</i>	<i>3008</i>	<i>1794</i>	<i>1289</i>

Bandymų metu nustatyta, kad TIFF rinkmenos neproporcingai didelės palyginti su JPEG ir PNG rinkmenomis, PNG rinkmenos dauguma atvejų yra mažesnės už JPEG rinkmenas, ir jų transformavimo trukmė yra mažesnė. Transformuojant *Microsoft PowerPoint* pateikčių turinį į PNG formatą, rinkmenos yra didesnės už JPEG rinkmenas, ir transformavimas trunka ilgiau. Tą paaiškina faktas, kad pateikčių skaidrėse buvo rasta fotografinių vaizdų (pvz., pateikties fonas).

Bandymai buvo pakartoti, naudojant kitą el. dokumentų imtį (100 el. dokumentų, bendras pradinių duomenų dydis – 39,15 MB, 169 turinio rinkmenos) 300 dpi raiška. Apibendrinus gautus rezultatus, nustatyta, kad tendencijos išliko tokios pačios:

Palyginamasis kriterijus	JPEG (85%)	PNG
Transformavimo trukmė (min.)	23	23
Bendras rezultatų dydis (MB)	593,4	366,9
<i>Rezultato dydis (MB) 100 MB pradinių duomenų</i>	<i>1516</i>	<i>937</i>

Atlikus bandymus taip pat buvo pastebėta, kad esant 100 dpi raiškai JPEG rinkmenos yra mažesnės už PNG rinkmenas, o transformavimas į jas trunka trumpiau, tačiau papildomų bandymų metu (žr. 3.4.4 skyrių) nustatyta, kad 100 dpi raiška tekstas nėra kokybiškai perskaitomas.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

3.4.2. Galimybių peržiūrėti formatus be papildomos programinės įrangos bandymai

Buvo išbandytos galimybės peržiūrėti dokumento vaizdą be specialiosios papildomos programinės įrangos interneto naršyklėje:

Operacinė sistema	Interneto naršyklė	PDF/A	XPS	TIFF	JPEG	PNG
Windows XP SP3	Internet Explorer 7.0	NE	NE	Ne ¹	Taip	Taip
	Mozilla Firefox 3.0	NE	NE	Ne ¹	Taip	Taip
	Opera 9.6	NE	NE	Ne ¹	Taip	Taip
Windows Vista Business	Internet Explorer 7.0	NE	Taip	Ne ³	Taip	Taip
	Mozilla Firefox 3.0	NE	Ne ²	Ne ³	Taip	Taip
	Opera 9.6	NE	Ne ²	Ne ³	Taip	Taip
Linux Fedora 9	Mozilla Firefox 3.0	Ne ⁴	NE ^{4,5}	Ne ⁴	Taip	Taip
	Opera 9.6	Ne ⁴	NE ^{4,5}	Ne ⁴	Taip	Taip

¹ Rinkmena buvo atidaryta naudojant kompiuteryje buvusią programą *Windows Picture&Fax Viewer*.

² Rinkmena buvo atidaryta naudojant *Internet Explorer 7* naršyklę.

³ Rinkmena buvo atidaryta naudojant kompiuteryje buvusią programą *Windows PhotoGallery* (programa naudoja neproporcingai daug atminties ir procesoriaus resursų).

⁴ Rinkmena buvo atidaryta naudojant kompiuteryje buvusią programą *Okular*.

⁵ Programa *Okular* rodo iškraipytą XPS formato turinį.

Formatų turiniui parodyti reikalinga įdiegti šią programinę įrangą:

Operacinė sistema	PDF/A	XPS
Windows XP SP3	<i>Adobe Reader</i> iš www.adobe.com (nemokamai)	<i>XPS Viewer</i> iš www.microsoft.com (nemokamai)
Windows Vista Business	<i>Adobe Reader</i> iš www.adobe.com (nemokamai)	<i>Nereikia</i> – įdiegta standartiškai
Linux Fedora 9	<i>Nereikia</i> – įdiegta standartiškai	Komerčinės priemonės (pvz., <i>SANAtch</i>)

3.4.3. Vaizdų saugojimo skirtingų parametų JPEG ir PNG formatais bandymai

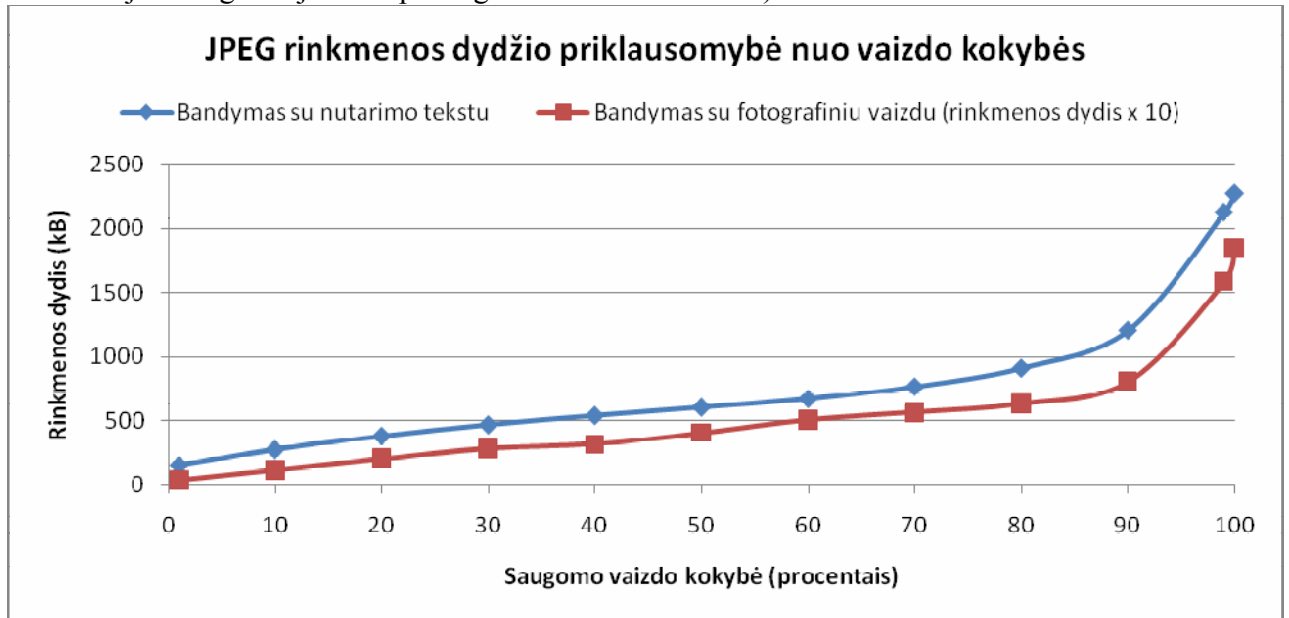
Buvo atlikti papildomi bandymai, siekiant nustatyti JPEG ir PNG paveikslėlių dydžio priklausomybę nuo glaudinimo parametru.

Bandymams naudotas 1 puslapio apimties Lietuvos Respublikos Vyriausybės 2006 m. gruodžio 7 d. nutarimas Nr. 1228 (žr. 1 priedą) DOC formatu (33 KB). Šis dokumentas, naudojant *Microsoft Office Word 2007* spausdinimo į *Microsoft Office Document Image Writer* savybę, išspausdintas į MDI formatą, o pastarasis išsaugotas 300 dpi raškos neglaudintu TIFF formatu (893 KB). Iš šio formato nemokamai platinama programa *IrfanView* (versija 3.99) buvo pagaminta 12 skirtingos kokybės (1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 99%, 100%) JPEG formato rinkmenų. Analogiškai buvo parengtos rinkmenos su fotografiniu vaizdu (2 priedas; TIFF rinkmenos dydis – 0,98 MB).

**Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas**

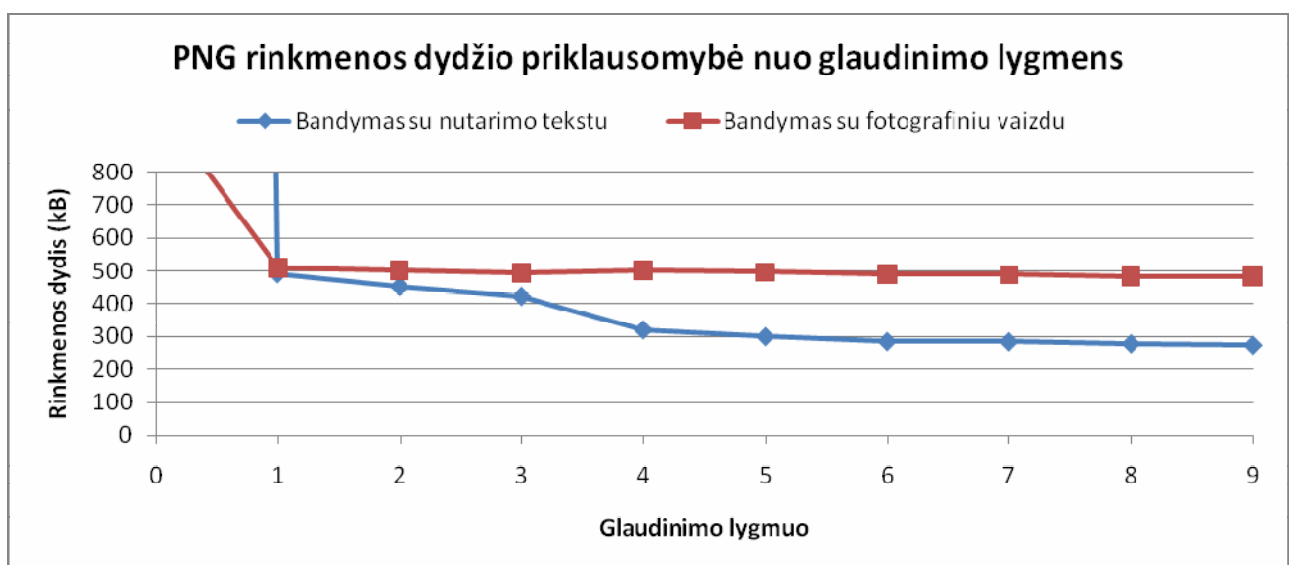
Rekomendacijos dėl elektroninio dokumento turinio formatų

Bandymo rezultatai apibendrinti diagramoje (fotografinių vaizdų rinkmenų dydis atvaizduojant diagramoje buvo padaugintas iš koeficiento 10):



Pastebima, kad vaizdo kokybę palaipsniui mažinant nuo 100% iki 70% JPEG rinkmenos dydis mažėja kur kas sparčiau, negu toliau mažinant kokybę iki 0%. Optimaliausias saugomo vaizdo kokybės ir rinkmenos dydžio taškas (balno taškas) yra ties 80% saugomo vaizdo kokybės riba.

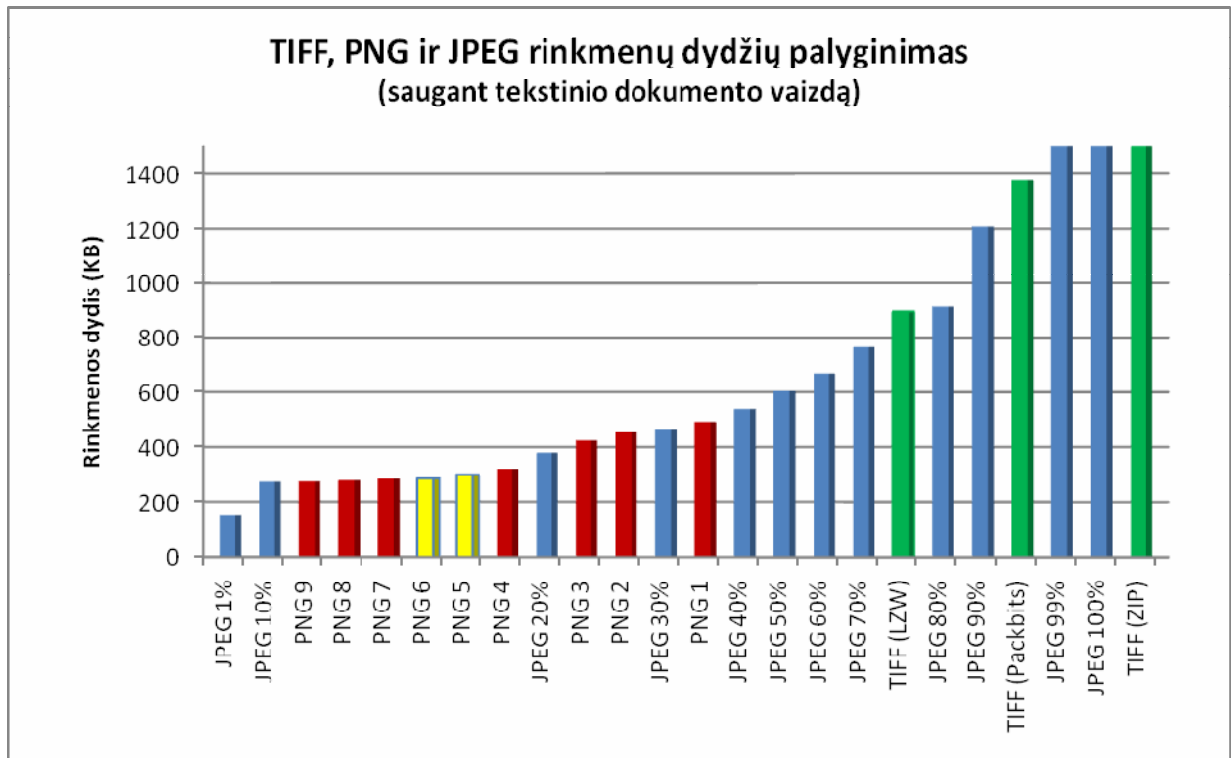
Analogiškai buvo pagaminta po 9 skirtingo glaudinimo lygmens (nuo 0 (neglaudinta) iki 9 (maksimalus glaudinimas)) PNG formato rinkmenos. Bandymo rezultatai apibendrinti diagramoje:



Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų

Pastebima, kad saugant dokumento teksto vaizdą (mėlyna linija) optimaliausius rezultatus teikia PNG vaizdų glaudinimo algoritmas, naudojant parametrus 5 arba 6. Saugant fotografinį vaizdą (ruda linija), rinkmenos dydis, didinant glaudinimo režimą, kinta labai nežymiai.

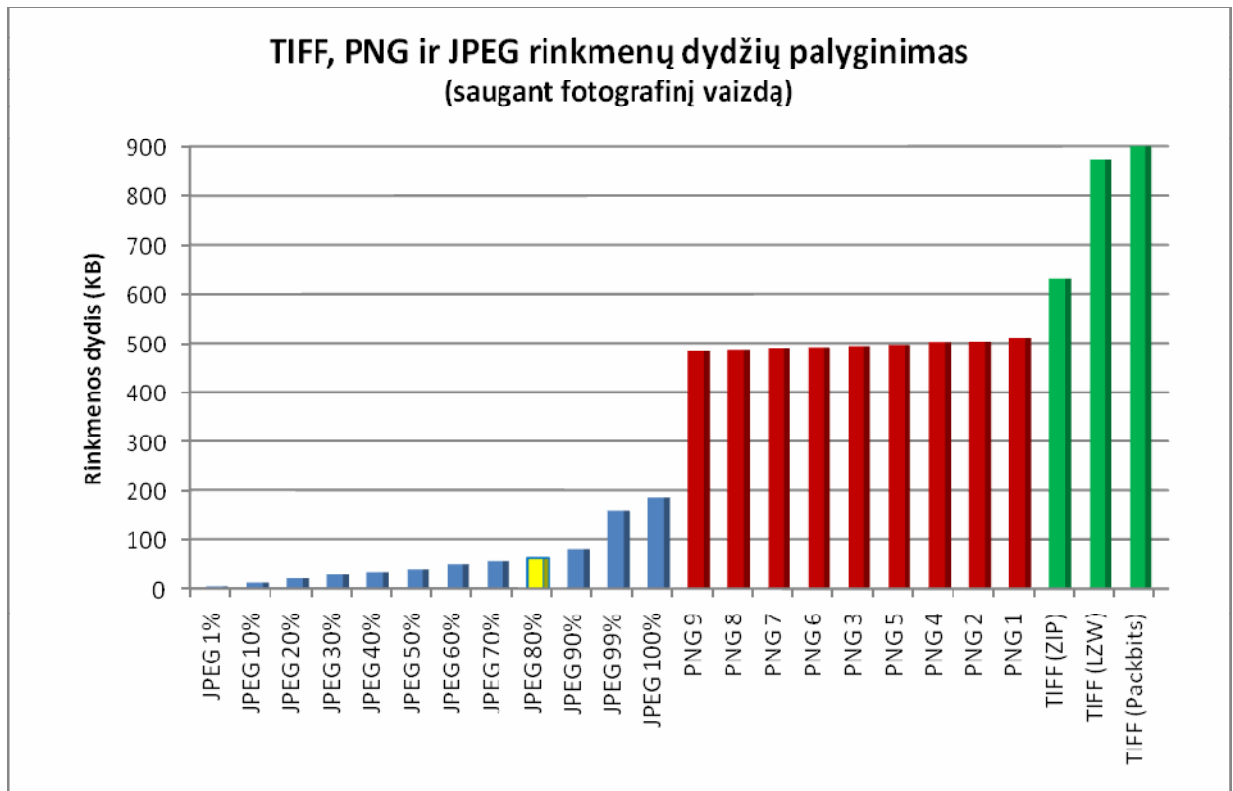


Lyginant PNG ir JPEG formatų rinkmenas, kuriose išsaugotas tekstinio dokumento vaizdas, matyti, kad PNG formato rinkmenų, kai vaizdai suglaudunami režimais 4-9, dydžiai yra mažesni už JPEG vaizdo 20% kokybės rinkmeną. Atkreiptinas dėmesys į tai, kad PNG formatas glaudindamas nepraranda vaizdinės informacijos, o JPEG naudoja informaciją prarandančius glaudinimo algoritmus.

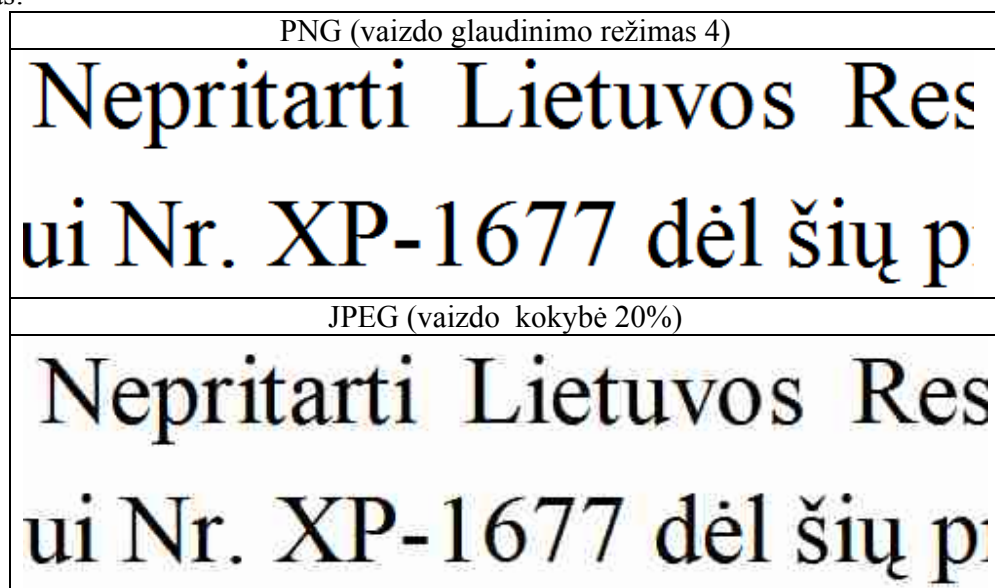
Saugant fotografinį vaizdą situacija yra atvirkščia: visos JPEG formato rinkmenos yra mažesnio dydžio už PNG ir TIF formatų rinkmenas. Optimaliausias vaizdo ir kokybės santykis gaunamas, JPEG formatu saugant 80% kokybės vaizdą:

Elektroninių dokumentų saugojimo informacinės sistemos prototipo
(bandomojo projekto) sukūrimas

Rekomendacijos dėl elektroninio dokumento turinio formatų



10 paveiksle pateiktas PNG formato rinkmenoje, kurioje 300 dpi tekstinio dokumento vaizdas saugomas suglaudintas 4 režimu, ir JPEG formato rinkmenoje (kokybė 20%) saugomų vaizdų palyginimas:



10 paveikslas. Vaizdo PNG ir JPEG (20%) formatais palyginimas.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Taip pat nustatyta, kad PNG formato rinkmenoje ir JPEG formato rinkmenoje (kokybė 50%) saugomi fotografiniai vaizdai vizualiai beveik nesiskiria.

3.4.4. Vaizdų saugojimo skirtinga raiška bandymai

Buvo atlikti papildomi JPEG ir PNG paveikslėlių naudojamumo bandymai, siekiant nustatyti raišką, reikalingą teksto kokybiškam perskaitymui. Bandymai buvo atliekami su 12 dydžio tekstu (1 priedas) ir su 5 dydžio tekstu bei fotografija (3 priedas).

Bandymo su 12 dydžio (1 priedas) tekstu rezultatai:

Raiška	Rinkmenos dydis (KB)			Teksto perskaitymas
	TIFF (LZW)	PNG (6)	JPEG (80%)	
300x300 dpi	893	285	911	Kokybiškai perskaitomas
250x250 dpi	687	258	710	Kokybiškai perskaitomas
200x200 dpi	506	189	513	Gera perskaitomas
150x150 dpi	272	126	343	Gera perskaitomas
100x100 dpi	133	74,5	193	Nekokybiškas
50x50 dpi	38,7	25,1	65,2	Neįskaitomas
PDF/A-1 formato rinkmenos dydis – 199 KB, tekstas perskaitomas kokybiškiausiai				

Bandymo su 5 dydžio tekstu ir fotografija (3 priedas) rezultatai:

Raiška	Rinkmenos dydis (KB)			Teksto perskaitymas
	TIFF (LZW)	PNG (6)	JPEG (80%)	
300x300 dpi	4140	2310	1480	Kokybiškai perskaitomas
250x250 dpi	2950	1680	1150	Gera skaitomas
200x200 dpi	1960	1130	835	Nekokybiškas
150x150 dpi	1070	711	541	Nekokybiškas
100x100 dpi	506	344	289	Neįskaitomas
50x50 dpi	135	100	96,3	Neįskaitomas
PDF/A-1 formato rinkmenos dydis – 330 KB, tekstas perskaitomas kokybiškiausiai				

Pastaba. Fotografinis vaizdas kokybiškai peržiūrimas esant 100 dpi raiškai ir 80% kokybei.

3.5. Formatų palyginimas

Lentelėje pateiktas TIFF, JPEG, PNG formatų palyginimas:

Kriterijus	TIFF	JPEG	PNG
Saugomo vaizdo tipas	Taškinė grafika	Taškinė grafika	Taškinė grafika
Formatas sukurtas	1986 m.	1992 m.	1996 m.
Saugomos spalvos	RGB - 24 ar 48 bitų, nespaltoto vaizdo - 8 ar 16 bitų,	RGB - 24 bitų, nespaltoto vaizdo - 8 bitų	RGB - 24 ar 48 bitų, nespaltoto vaizdo - 8 ar 16 bitų,

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Kriterijus	TIFF	JPEG	PNG
	spalvų paletė – nuo 1 iki 8 bitų, vienspalvis - 1 bitas		spalvų paletė – nuo 1 iki 8 bitų, vienspalvis - 1 bitas
Vaizdo glaudinimas	Informacijos neprarandantis, paprastai nenaudojamas	Informacijos neprarandantis arba informaciją prarandantis pagal vaizdo kokybės procentą	Informacijos neprarandantis, 10 lygių
Glaudavimo algoritmai	Neglaudinama; modifikuotas Huffmano; PackBits; CCITT T.4; CCITT T.6; LZW; JPEG; ZIP	Tikimybinio spėjimo (100%) arba DCT (1-99%) su Huffmano arba aritmetiniu glaudinimu	Vaizdo filtravimas + deflate (zlib)
Galimybė rinkmenoje saugoti keletą vaizdų	Yra	Nėra	Nėra
2008-12-03 dieną išbandytų 40 ADOC specifikaciją atitinkančių dokumentų grupės (bendras dydis – 23 MB) transformuoto turinio dydis (300 dpi)	1 GB (ZIP)	1 GB (80%)	518 MB
Išbandyto 33 KB dydžio DOC dokumente saugomo tekstinės informacijos 1 puslapio vaizdo dydis (2480x3508, 300 dpi, 16,7 mln. spalvų gama)	893 KB	379 KB (20%) – 911 KB (80%)	278 KB (8 lygmuo) – 453 KB (2 lygmuo)
Išbandytos fotonuotraukos dydis (720x480, 73 dpi, 16,7 mln. spalvų gama)	980 KB	20,3 KB (20%) – 63,1 KB (80%)	484 KB (8 lygmuo) – 503 KB (2 lygmuo)
Standartai	ISO 12234-2:2001 (TIFF/EP); ISO 12639:2004 (TIFF/IT)	ISO/IEC 10918-1:1994	RFC 2083 (1996); ISO/IEC 15948:2003; ISO/IEC 15948:2004
Panašūs formatai, atmainos	Komplikuotas formatas: labai daug atmainų	Keletas atmainų: JIF, JFIF, Exif, JPEG2000	Su animacijomis: MNG, APNG, ANG, aNIM/mPNG

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

3.6. Interneto bendruomenės patirtis

Šaltinis: <http://paleo.org/BKsamples/index.html>. „Vaizdo formatų palyginimas“, Edward Reid.

„Atkreiptinas dėmesys, kad aš nelyginu TIFF formatų su kitais vaizdų formatais. TIFF formatas yra labiau vaizdo rinkmenų formatas nei vaizdo formatas. Yra galimybė TIFF rinkmenoje saugoti skirtingo formato vaizdus, todėl TIFF būtų sudėtinga lyginti su kitais vaizdo formatais. TIFF tapo vaizdinę informaciją neprarandančių formatų, ypač nesuspausto vaizdo formatų, akronimu. Šiuo metu (2005 m. balandį), kai PNG formatas yra plačiai naudojamas šiuolaikinėje grafinėje programinėje įrangoje, liko labai mažai priežasčių rinktis TIFF formatą.

Rekomendacijos pagal paskirtį:

Fotografinių vaizdų ilgalaikiam saugojimui ir vėlesniam jų redagavimui: PNG arba aukštos kokybės JPEG formatas.

Fotografinių vaizdų platinimui internete: vidutinės kokybės JPEG formatas, mažos kokybės JPEG formatas.

Vaide esančių tekstų atpažinimui (OCR): GIF (300 dpi), geresnei kokybei – GIF (600 dpi).

Vienspalvių (juodo teksto baltame fone) aukštos kokybės dokumentų vaizdų archyviniam saugojimui: GIF, mažiausia skenuojama raiška 600 dpi.

Dokumentų vaizdų archyviniam saugojimui: PNG, mažiausia skenuojama raiška 300 dpi.

PNG formatas (300dpi): geriausias visų (išskyrus vienspalvių) vaizdų, kurie vėliau gali būti redaguojami, saugojimui. Geriausias archyviniam saugojimui. Tinkamas visiems spalvotiems vaizdams saugoti, įskaitant ir dokumentų vaizdus, kai svarbu išsaugoti jo tikrąjį vaizdą.

JPEG formatas (300dpi, žema kokybė (1-50%)): tinkamas fotografinių vaizdų platinimui internete, kadangi tokie vaizdai išlieka kokybiški ir užima nedaug vietos. Nenaudokite archyvinėms kopijoms ar kopijoms, kurias vėliau reikia redaguoti, gaminti. Nenaudokite dokumentų vaizdams saugoti.

JPEG formatas (300dpi, vidutinė kokybė (50-80%)): geriausias fotografinių vaizdų platinimui internete. Rinkmenos dydis yra tik šiek tiek didesnis nei žemos kokybės JPEG formato, tačiau vaizdo kokybė nepalyginamai geresnė. Nenaudokite archyvinėms kopijoms ar kopijoms, kurias vėliau reikia redaguoti, gaminti. Nenaudokite dokumentų vaizdams saugoti.

JPEG formatas (300dpi, aukšta kokybė (80-100%)): labai tinkamas fotonuotraukų, kurių nereikia redaguoti arba tik minimaliai poredaguoti, archyviniam saugojimui. Dėl didelės rinkmenų apimties mažai tinkamas fotografinių vaizdų platinimui internete. Iš esmės, formatas tinkamas dokumentų vaizdų saugojimui, jeigu vienspalvio vaizdo nepakanka, ir neturite pakankamai vietos PNG formato vaizdams saugoti.

GIF formatas (600dpi): geriausias vienspalvių (juodas tekstas baltame fone) dokumentų vaizdams saugoti ir teksto optiniam atpažinimui (OCR).“

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Šaltinis: http://en.wikipedia.org/wiki/Portable_Network_Graphics. „Portable Network Graphics“, Vikipedija.

„PNG lyginant su JPEG:

Saugant fotografinius vaizdus, JPEG formato rinkmenos yra mažesnės nei PNG rinkmenos. Lyginant su aukštos kokybės JPEG formato rinkmenomis, PNG rinkmenos, jose saugant fotografinius vaizdus, paprastai yra 5-10 kartų didesnės, o kokybės skirtumas tarp skirtingų formatų beveik nepastebimas.

PNG formatas yra geresnis nei JPEG, kai reikia saugoti vaizdus, kuriuos sudaro tekstas, brėžiniai ar kita didelio kontrasto vaizdinė informacija.

JPEG netinkamas vaizdams, kuriuos vėliau numatyta redaguoti, saugoti, kadangi JPEG naudoja informaciją prarandantį glaudinimo algoritmą.

PNG lyginant su GIF:

Paprastai gerai optimizuotos ir suglaudintos PNG formato rinkmenos yra 10-50% mažesnės nei GIF formato rinkmenos; PNG formatas numato fono permatomumo savybę; leidžia saugoti vaizdus tikromis spalvomis (maksimalus GIF formate numatytas spalvų skaičius – 256). GIF formatas leidžia animaciją, tuo tarpu PNG formatas yra statinis.“

Šaltinis: <http://home.tiscali.nl/developerscorner/imaging/image-saving.htm>. „Vaizdų saugojimas“, Francky Kleyneman, 2006 m.

Vaizdo tipas	GIF	PNG	JPEG (99%)	JPEG (75%)	JPEG (50%)
25 spalvų kvadratėliai	2,3 kB	0,8 kB	7,6 kB	-	3,5 kB (prasta kokybė)
117 spalvų fonas	3,5 kB	0,78 kB	1,8 kB	-	0,85 kB
11185 spalvų fotonuotrauka	11 kB (254 spalvos)	34 kB	-	5 kB	4 kB (prasta kokybė)
720x480px dydžio fotonuotrauka	197 kB	564 kB	150 kB	58 kB	-

3.7. Išvados

TIFF formatas yra vienas pirmųjų plačiausiai naudotų taškinės grafikos vaizdų saugojimo formatų, sukurtas, kai dar nebuvo sukurtų efektyvių vaizdo glaudinimo algoritmų. Nors formatas jais buvo papildytas vėliau, TIFF tapo neglaudinto vaizdo ir didelės apimties rinkmenų akronimu. Nuo 1996 metų, kai buvo sukurtas informacijos neprarandantis PNG formatas, vaizdą išsaugantis mažiausiai 2 kartus mažesnėse rinkmenose, priešasčių vaizdų saugojimui rinktis TIFF formatą liko labai nedaug, daugiausia dėl poreikio užtikrinti turimų programų ar sistemų sąveikumą.

Tačiau TIFF formatas išsiskiria galimybe vienoje rinkmenoje saugoti keletą vaizdų, o tai padeda efektyviau išsaugoti iš kelių puslapių susidedančio dokumento turinio atvaizdo puslapių struktūrą.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

Tačiau lyginant TIFF formatų rinkmenas (naudojant standartuose nedokumentuotą LZW glaudinimo algoritmą) su glaudintų vaizdų JPEG ir PNG rinkmenomis, TIFF rinkmenos yra neproporcingai didelės, todėl nėra patogios informacijos keitimuisi internete. Dėl šios priežasties nėra viena iš išbandytų interneto naršyklių (*Internet Explorer 7, Mozilla Firefox 3, Opera 9.6*) šio formato rinkmenų nerodo be kompiuteryje įdiegtų papildomos programinės įrangos. Atkreipiamas dėmesys į tai, kad *Internet Explorer 3* versija, išleista 1995 metais kartu su *Windows 95* operacine sistema, turėjo galimybę parodyti TIFF formato rinkmenas. Vėlesnėse naršyklės versijose ši savybė buvo panaikinta.

Atsižvelgiant į tai, darytina išvada, kad TIFF formatas, lyginant jį su JPEG ir PNG formatais, yra mažiausiai tinkamas naudoti elektroninio dokumento vaizdo peržiūrai internete.

Iš JPEG formato analizės matyti, kad šis formatas buvo specialiai sukurtas fotografiniams vaizdams saugoti, ir bandymų rezultatai parodė, kad geresnio formato už JPEG, lyginant su TIFF ir PNG formatais, šioje srityje nėra.

Tačiau šiame projekte išbandomos pagrindinės oficialių elektroninių dokumentų perėmimo ir ilgalaikio bei nuolatinio saugojimo procedūros. Viena iš procedūrų – elektroninio dokumento turinio transformavimas į peržiūrai internete tinkamus formatus. Siekiama išsaugoti dokumentų, kurių turinį sudaro daugiausia tekstinė informacija, atvaizdą. Tekstinės informacijos vaizdas yra kontrastinis, susidedantis daugiausia iš baltame fone esančių juodos spalvos simbolių, linijų, taškų ir kreivių. JPEG formate naudojamas vaizdo glaudinimo algoritmas, parenkantis taško spalvą, atsižvelgiant į gretimus taškus, veikia ne visai korektiškai, vaizdinę informaciją glaudina neefektyviai.

Bandymų rezultatai rodo, kad PNG rinkmenos tekstinių dokumentų atvaizdui saugoti yra mažesnės už JPEG rinkmenas, nepaisant to, kad PNG rinkmenose saugomas informacijos neprarandantis vaizdas.

Atsižvelgiant į tai, darytina išvada, kad elektroninio dokumento turinio atvaizdo išsaugojimui tinkamiausias yra PNG formatas, tačiau fotografinių vaizdų saugojimui reikėtų rinktis JPEG formatą.

Internetu taip pat yra plačiai naudojamas GIF formatas. Tačiau šis formatas nėra tarptautinių organizacijų standartizuotas dėl labai aiškios priežasties: korporacija „Unisys“ pareiškė patento teises į GIF formate naudojamą LZW glaudinimo algoritmą. 2004 metais šių patentinių teisių galiojimas baigėsi. Per patento galiojimo laikotarpį buvo sukurtas PNG formatas (sertifikuotas ISO) ir dar keletas formatų, kurie yra pažangesni už GIF formatą. Dėl šių priežasčių darytina išvada, kad GIF formatas nėra tinkamas naudoti elektroninio dokumento vaizdo peržiūrai internete.

Atlikus vaizdų saugojimo skirtinga raiška bandymus, įsitikinta, kad kokybiškiausiai tekstą, išsaugotą taškinės grafikos vaizdu, perskaityti galima esant 300 dpi raiškai. Tekstas pakankamai gerai skaitomas esant ir 250 dpi raiškai. Fotografinis vaizdas kokybiškai peržiūrimas esant 100 dpi raiškai.

Bandyto metu taip pat nustatyta, kad geriausiai dokumento atvaizdas peržiūrimas naudojant PDF/A-1 rinkmenos formatą ir *Adobe Reader* programines priemones. Minėto formato rinkmenos dydis yra mažiausias, lyginant su TIFF, JPEG ir PNG taškinės grafikos vaizdų rinkmenomis, kurių turinį galima kokybiškai peržiūrėti (net 1,5 – 3 kartus), teikiamomis peržiūros priemonėmis dokumento vaizdą yra patogiau peržiūrėti ir spausdinti – reikia ilgai derinti interneto naršyklę, kai norima kokybiškai išspausdinti 300 dpi dydžio paveikslėlį. Atsižvelgiant į tai, siūloma atsisakyti idėjos elektroninio dokumento atvaizdą transformuoti į taškinės grafikos vaizdo saugojimo formatus (TIFF, JPEG, PNG, GIF), kadangi vektorinės grafikos priemonėmis išsaugotas vaizdas, naudojamas PDF/A-1 formatu (taip pat ir XPS formatu), teikia kokybiškiausią dokumento turinio vaizdą ir ši kokybė nepriklauso nuo raiškos, kuri naudojama dokumento turiniui peržiūrėti.

UAB „Sintagma“, UAB „MitSoft“, UAB „Etnomedijos Intercentras“	VX-2008/08-LAD1/F8-13
Elektroninių dokumentų saugojimo informacinės sistemos prototipo (bandomojo projekto) sukūrimas	
Rekomendacijos dėl elektroninio dokumento turinio formatų	

4. Rekomendacijos

Atlikę formatų analizę ir palyginimą, sukūrę elektroninio dokumento turinio transformavimo į ilgalaikio saugojimo formatus priemonę ir elektroninio dokumento turinio transformavimo į peržiūrai internete tinkamus formatus priemonę, ir šias priemones išbandę, rekomenduojame:

1. Elektroninio dokumento turinio ilgalaikiam ir nuolatiniam saugojimui naudoti ISO 19005-1:2005 standartą atitinkantį PDF/A-1 formatą. Jeigu turinio transformavimo programinė priemonė leidžia pasirinkti suderinamumo su standartu lygmenį (PDF/A-1a ar PDF/A-1b), rekomenduotinas PDF/A-1a suderinamumo su standartu lygmuo.
2. Elektroninio dokumento turinio transformavimui į peržiūrai internete tinkamą taškinės grafikos vaizdo formatą naudoti ISO/IEC 15948:2004 standartą atitinkantį PNG formatą. Rekomenduojama vaizdus saugoti ne mažesne nei 300 dpi raiška, naudoti glaudinimo algoritmus – 5 arba 6.
3. Elektroninių fotografinių vaizdų (jei tokių bus) archyviniam saugojimui ir peržiūrai internete naudoti ISO/IEC 10918-1:1994 standartą atitinkantį JPEG formatą. Rekomenduojama vaizdus saugoti ne mažesne nei 100 dpi raiška ir 80% glaudinimo kokybe.
4. Atsisakyti idėjos oficialių elektroninių dokumentų turinį transformuoti į taškinės grafikos vaizdo saugojimo formatus. Elektroninio dokumento turinį rekomenduojame transformuoti tik į PDF/A-1 formatą, skirtą naudoti ir elektroninio dokumento turinio ilgalaikiam bei nuolatiniam saugojimui, ir peržiūrai internete.
5. Atsižvelgiant į sparčią IT technologijų raidą, sprendimus dėl tinkamiausių formatų parinkimo reguliariai (kas 3-5 metus, o esant poreikiui – dažniau arba rečiau) peržiūrėti.



LIETUVOS RESPUBLIKOS VYRIAUSYBĖ

NUTARIMAS DĖL LIETUVOS RESPUBLIKOS AKCIZŲ ĮSTATYMO 35 STRAIPSNIO PAPILDYMO ĮSTATYMO PROJEKTO NR. XP-1677

2006 m. gruodžio 7 d. Nr. 1228
Vilnius

Vadovaudamasi Lietuvos Respublikos Seimo statuto (Žin., 1994, Nr. [15-249](#); 1999, Nr. [5-97](#); 2000, Nr. [86-2617](#); 2004, Nr. [165-6025](#)) 138 straipsnio 3 dalimi ir atsižvelgdama į Lietuvos Respublikos Seimo valdybos 2006 m. spalio 6 d. sprendimo Nr. 1175 2 punktą, Lietuvos Respublikos Vyriausybė n u t a r i a :

Nepritarti Lietuvos Respublikos akcizų įstatymo 35 straipsnio papildymo įstatymo projektui Nr. XP-1677 dėl šių priežasčių:

1. 2003 m. spalio 27 d. Tarybos direktyva 2003/96/EB, pakeičianti Bendrijos energetikos produktų ir elektros energijos mokesčių struktūrą (toliau vadinama – Tarybos direktyva 2003/96/EB), reglamentuojanti energetikos produktų ir elektros energijos apmokestinimo akcizais tvarką Europos Sąjungoje, nenumato galimybės nuo akcizų atleisti dyzelinius degalus, kuriuos naudoja visas vietinis viešasis keleivinis transportas (įskaitant maršrutinius taksi), todėl taikyti tokią akcizų lengvatą būtų galima tik tuo atveju, jeigu Europos Sąjungos Taryba suteiktų Lietuvai išimtį ir Tarybos direktyvos 2003/96/EB III priede būtų numatyta Lietuvai teisė taikyti akcizų lengvatą dyzeliniams degalams, kuriuos naudoja vietinis viešasis keleivinis transportas (įskaitant maršrutinius taksi).

2. Jeigu nuo 2007 m. liepos 1 d. dyzeliniai degalai, skirti naudoti vietiniam viešajam keleiviniam transportui (įskaitant maršrutinius taksi), būtų neapmokestinami akcizais, 2007 metais Lietuvos Respublikos valstybės biudžetas netektų apie 45 mln. litų pajamų (2008 metais – apie 101 mln. litų). Lietuvos Respublikos Vyriausybė 2006 m. spalio 11 d. nutarimu Nr. 988 „Dėl Lietuvos Respublikos 2007 metų valstybės biudžeto ir savivaldybių biudžetų finansinių rodiklių patvirtinimo įstatymo projekto ir su juo susijusių teisės aktų projektų pateikimo Lietuvos Respublikos Seimui“ (Žin., 2006, Nr. [110-4172](#)) jau pritarė Lietuvos Respublikos 2007 metų valstybės biudžeto ir savivaldybių biudžetų finansinių rodiklių patvirtinimo įstatymo projektui, todėl nusprendus neapmokestinėti akcizais dyzelinių degalų, skirtų naudoti vietiniam viešajam keleiviniam transportui (įskaitant maršrutinius taksi), reikėtų atitinkama suma mažinti šiame įstatymo projekte asignavimų valdytojams numatytus biudžeto asignavimus.

Ministras Pirmininkas

Gediminas Kirkilas

Finansų ministras

Zigmantas Balčytis

Fotografija

Dydis 25.1 x 16.7 cm, 720 x 480 taškų;
Raiška 73 x 73 dpi,
Maksimalus spalvų skaičius – 16,7 mln.;
108046 skirtingos spalvos.



EUROPOS BENDRIJA					A IŠSIUNTIMO/EKSPORTO ĮSTAIGA						
2	2 Siuntėjas/Eksportuotojas Nr.				1 DEKLARACIJA						
	3 Lapai				4 Krov. aprašai						
	5 Iš viso prekių				6 Iš viso vietų		7 Registracijos numeris				
	8 Gavėjas Nr.				9 Asmuo, atsakingas už finansinį atsiskaitymą Nr.						
	10 Pirmoji šalis		11 Prekiaujančioji		13 BŽŪP						
	gavėja		šalis								
	14 Deklarantas/Atstovas Nr.				15 Šalis siuntėja/eksportuotoja		15 Šalies siunt./eksp. kodas		17 Šalies		
					a		b		gavėjos kodas		
					16 Kilmės šalis		17 Šalis gavėja				
	18 Transporto priemonė ir jos registracijos šalis išvykstant				19 Kont.		20 Pristatymo sąlygos				
21 Aktyvioji transporto priemonė ir jos registracijos šalis vykstant per sieną				22 Valiuta ir bendroji faktūrinė vertė			23 Kursas		24 Sandorio rūšis		
25 Transporto rūšis pasienyje		26 Vidaus transporto rūšis		27 Pakrovimo vieta		28 Finansiniai ir banko duomenys					
29 Išvykimo per sieną įstaiga				30 Prekių buvimo vieta							
31 Krovinio vietos ir prekių aprašymas	Ženkliai ir numeriai. Konteinerių Nr. Skaičius ir rūšis				32 Prekės Nr.		33 Prekių kodas				
	34 Kilmės šalies kodas		35 Bruto masė (kg)		a		b				
	37 PROCEDŪRA		38 Neto masė (kg)		39 Kvota						
	40 Bendroji deklaracija/pirminis dokumentas										
	41 Papildomas matavimo vienetas				PIK						
44 Papildoma informacija/ pateikiami dokumentai/serifikatai ir leidimai				46 Statistinė vertė							
47 Mokesčių apskaičiavimas	Tipas	Mokesčio pagrindas	Norma	Suma	MB	48 Mokejimo atidėjimas		49 Muitinės sandėlis			
	B APSKAIČIAVIMO DETALIZACIJA										
Iš viso											
50 Vykdytojas Nr.				Parašas:		C IŠVYKIMO ĮSTAIGA					
51 Numatomas tranzito įstaigos (ir šalys)				atstovaujamas Vieta ir data.							
52 Garantija negalioja				Kodas		53 Paskirties įstaiga (ir šalis)					
D IŠVYKIMO ĮSTAIGOS TIKRINIMAS				Antspaudas:		54 Vieta ir data:					
Rezultatas:						Deklaranto/atstovo pavadinimas (vardas, pavardė) ir parašas:					
Uždėtų plombų skaičius: aprašymas:											
Terminas (data):											
Parašas:											

