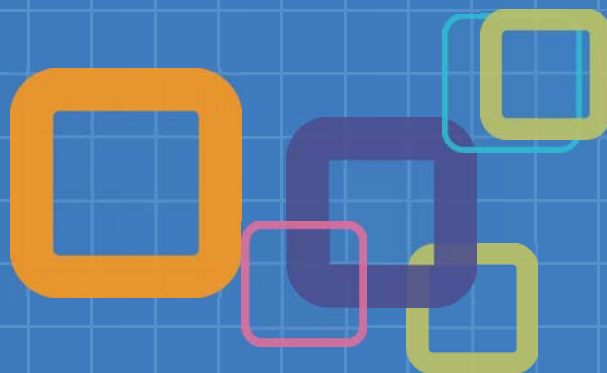




Standard OMPI ST. 22

fil. Adriana NEGOIȚĂ, OSIM;
fil. Adriana-Cătălina GRIGORE, OSIM

Oficiul de Stat pentru Invenții și Mărci este angrenat în proiectul Oficiului European de Brevete (EPO) denumit "Calitatea la sursă" (QaS). Acest proiect se referă atât la transmiterile curente de date publicate de Oficiile Naționale, cât și la cele anterioare, începând cu anul 1973, scopul urmărit fiind armonizarea și standardizarea datelor transmise, pentru furnizarea unor servicii mai bune către utilizatori (date ușor accesibile, care cuprind o perioadă mai mare).

În vederea eliminării problemelor identificate în documentele depuse la OSIM, privind cererile de brevet de invenție, și anume:

- dimensiunea prea mică a fontului;
- text scris pe fundal negru/închis la culoare;
- rezoluție mică la scanare/pagină scanată oblic/incomplet;
- caractere ilizibile;
- ștampilă amplasată peste text;
- utilizarea mai multor tipuri de fonturi;
- text scris de mână;
- text amestecat cu formule;
- formule chimice subliniate,

vă oferim traducerea Standardului 22 al Oficiului Internațional al Proprietății Intelectuale (WIPO), care cuprinde regulile ce trebuie urmate în formatarea cererilor de brevet depuse pe hârtie sau electronic.

STANDARDUL ST.22/OMPI - (traducere)

RECOMANDARE PENTRU FORMATAREA CERERILOR DE BREVET ÎN VEDEREA FACILITĂRII RECUNOAȘTERII OPTICE A CARACTERELOR (OCR)

Revizuire adoptată de Grupul de elaborare a documentației privind standardele SCIT la cea de-a X-a sesiune, din 21 noiembrie 2008

INTRODUCERE

1. Această Recomandare se aplică la cererile de brevet depuse pe hârtie sau în format electronic (e-filed), dar care au textul cererii depus sub formă de imagine (de exemplu, imagini PDF sau TIFF).

2. Această Recomandare a fost întocmită pentru a ajuta la pregătirea de cereri de brevet într-o formă tipărită, potrivită pentru producția ulterioară a unei arhive electronice, digitale, a conținuturilor cererilor de brevet, prin intermediul programelor de Recunoaștere Optică a Caracterelor (OCR).

3. Această Recomandare a fost întocmită pe baza experiențelor mai multor oficii la utilizarea programelor de OCR-are. A fost redactată cu scopul de a avea cea mai mică rată de eroare posibilă la stadiul de citire automată a textului cererilor de brevet, dar permițând, în același timp, și citirea personală eficientă a documentului. Se va lua în considerare că acest document nu oferă recomandări detaliate pentru limbile japoneză și coreeană; numărul de depuneri de texte în format electronic, din numărul total de depuneri anuale, reprezintă peste 90% în aceste țări, ceea ce înseamnă că acest Standard nu se aplică în practică celor două limbi: japoneză și coreeană.



4. Scopul principal al producerii unei arhive digitale de cereri de brevet este acela de a permite publicarea cu ușurință a acelei cereri, într-un format compus, folosind tehnici de tipărire computerizate, și de a îmbunătăți astfel prezentarea și valoarea documentelor de brevet, în avantajul tuturor beneficiarilor. Un alt scop este acela de a crea o bază de date în care poate fi citit întregul text al unui document publicat, astfel încât să se poată beneficia mai târziu de posibilitățile oferite de căutare completă de text pe calculator.

DEFINIȚII

5. În cadrul acestei Recomandări, expresia “cerere de brevet” se referă la cereri de brevete de invenție, brevete de invenție, cereri de modele de utilitate, modele de utilitate, cereri de certificate sau certificate suplimentare de protecție.

6. Se spune că o formulă matematică sau chimică este “complexă” dacă nu poate fi exprimată ca o înșiruire liniară de caractere, fiecare caracter având un atribut de subscript/superscript. O formulă este complexă în special când conține o serie de attribute subscript/superscript sau, eventual, simbolurile matematice pentru sumă, interval sau produs.

7. O casetă delimitată a unui caracter/set de caractere este cel mai mic dreptunghi aliniat cu axa, care include toate părțile caracterului/setului de caractere.

8. Termenul “cursiv” se referă la o formă stilizată a scrisului de mână, unde toate literele din cuvinte sunt conectate, făcând ca un cuvânt să fie o singură linie complexă. Despre fonturi se spune că sunt cursive dacă sunt concepute în așa fel încât să imite scrisul de mână.

CREAREA ORIGINALULUI

9. O cerere de brevet este deseori pregătită cu ajutorul programelor de procesare de text. Experiențele anterioare au arătat că cel mai eficient format care urmează a fi utilizat, și care ar permite utilizarea în mod fiabil a programului OCR, este cel definit în Standardul 1073/II al Organizației internaționale de standardizare (ISO), așa-numitul format OCR-B.

SUPORT DE HÂRTIE DACĂ DOCUMENTUL SE DEPUNE PE HÂRTIE

10. Pentru a facilita scanarea, suportul de hârtie al aplicației tipărite ar trebui să aibă următoarele caracteristici:

- (a) Hârtia trebuie să fie rezistentă, albă și curată.
- (b) Greutatea hârtiei trebuie să fie între 70, preferabil 80, și 120 g/m².
- (c) Mărimea hârtiei trebuie să fie, de preferat, A4, 201 mm x 297 mm sau 8½ pe 11 inci (care este standardul nord-american *de facto*).
- (d) Foile trebuie să nu prezinte cute sau găuri, și nu trebuie să fie laminate.
- (e) Hârtia nu trebuie să fie absorbantă, pentru a se evita impregnarea cu tuș (de exemplu, când se folosește o imprimată cu ink jet).

RECOMANDĂRI PENTRU CONFIGURAȚIA PAGINII

11. Imprimarea pe două fețe ar trebui evitată. Dacă acest lucru nu este posibil, trebuie să se folosească hârtie suficient de opacă pentru a asigura o scanare curată față/verso.

12. Caracterele trebuie să fie negre, pe un fundal alb.

13. O margine de minimum 2 centimetri trebuie să apară în părțile de sus, de jos și din dreapta ale fiecărei foi, și o margine de cel puțin 2,5 centimetri pe partea stângă. Toate referințele deponentului sau ale reprezentantului legal al acestuia trebuie să apară în marginea superioară (vezi Anexa 1).

14. Numerotarea rândurilor ar trebui evitată. Dacă este necesară, numerele de rând trebuie să fie scrise cu caractere arabe, în marginea din stânga, la cel puțin 1 cm în afara marginilor textului, după cum se arată în Anexa 1. Mărimea fontului numerelor de rând trebuie să fie de cel puțin 12.

15. Numerotarea paginilor trebuie să fie realizată cu caractere arabe, fără alte caractere delimitatoare. Numerele de pagină trebuie, preferabil, să fie centrate, la începutul sau la baza foii, în margine, după cum se arată în Anexa 1.

16. Descrierea, revendicările și rezumatul trebuie să înceapă fiecare pe o pagină nouă. În plus, primul cuvânt imprimat pe prima pagină a fiecăreia dintre cele trei părți ale cererii, menționate anterior, trebuie să specifice partea corespondentă (în limbajul fiecărei cereri); paragraful de revendicări trebuie să fie numerotat secvențial. Formatul numerotării revendicărilor trebuie să lase posibilitatea unei separări clare între numărul revendicării și textul fiecărei revendicări. Formatele recomandate sunt ori folosirea numeralelor arabe, urmate de punct, ori

folosirea cuvântului “Revendicare” - sau echivalentul acestuia, în limbajul cererii de brevet - urmat de un spațiu și de numărul arab al revendicării, textul următor al revendicării fiind aliniat la dreapta față de numărul revendicării, la cel puțin 1 cm în ambele cazuri.

17. Paginile trebuie să fie constituite din paragrafe cu o singură coloană (paragrafe de text sau paragrafe care includ o imagine încorporată).

18. Paginile care conțin paragrafe trebuie să aibă orientarea tip portret.

19. Orientarea tip vedere este de evitat. Aceasta este acceptată doar în cazul paginilor care conțin figuri sau tabele încorporate, care nu ar avea loc într-o orientare tip portret.

20. Fiecare pagină trebuie să conțină text orientat într-o singură direcție.

21. Paginile tip vedere trebuie să fie întoarse la 90 de grade în sens contrar acelor de ceasornic, pentru buna integrare în setul de pagini tip portret.

22. Se recomandă evitarea folosirii notelor de subsol, textelor de pe margine sau antetelor, cu excepția celor indicate în paragraful 14 (numerotarea rândurilor), paragraful 15 (numerotarea paginilor) și a celor care includ referințe privind dosarul deponentului, scrise în colțul din stânga-sus al marginii.

RECOMANDĂRI PENTRU CONFIGURAȚIA PARAGRAFULUI

23. Se recomandă ca tabelele, formulele chimice complexe, formulele matematice complexe, imaginile și figurile să fie separate de paragrafele de text. Se recomandă ca elementele de acest fel să fie înconjurare de margini goale, deasupra și dedesubt, de cel puțin 1 cm, care să cuprindă lățimea paginii.

24. Imaginile și figurile trebuie să fie incluse în secțiunea de “Figuri” și să se facă referințe la ele în secțiunile “Descriere” și “Revendicări” ale cererii de brevet.

25. Imaginile și figurile trebuie să fie alb-negru (imaginile gri trebuie evitate, pentru că se pierd informații când sunt scanate sau convertite în alb-negru). Figurile trebuie să conțină linii clare, care să fie destul de groase pentru a fi reprezentate bine la o rezoluție de 300 dpi.

26. Paragrafele de text sau adnotările scrise de mână trebuie evitate. Dacă sunt necesare, acestea vor fi considerate imagini încorporate, și trebuie să urmeze recomandarea din paragraful 23.

27. Textul trebuie scris având spațierea dintre rânduri la un rând și jumătate.

28. Paragrafele trebuie să fie separate cu o spațiere care este de cel puțin două ori mai mare decât cea dintre rândurile din interiorul paragrafului.

29. Toate caracterele cuprinse într-un rând din paragraf trebuie să aibă baza aliniată cu grija, mai puțin în cazul caracterelor subscript și superscript, cum este indicat în paragraful 35.

30. Paragrafele de text aliniat trebuie evitate. Dacă se aplică, spațierea dintre cuvinte trebuie să fie cel puțin la fel de mare ca în cazul textului nealiniat. Textul aliniat poate să împiedice sistemele de OCR-are să identifice corect limitele cuvintelor dintr-un paragraf.

31. Când este posibil, scindarea cuvântului prin intermediul cratimei trebuie evitată (de exemplu, la sfârșit de rând sau în celulele unui tabel). Acest lucru nu se aplică în cazul limbilor care folosesc cuvinte compuse (de exemplu, limba germană).

RECOMANDĂRI PENTRU TABELE

32. Se folosește doar fundal alb.
33. Tabelele trebuie să aibă borduri. Bordurile trebuie să fie doar linii solide mai groase de 1,5 puncte.

RECOMANDĂRI DE FONT

34. Mărimea minimă de font recomandată este de 12, 14 fiind preferabil. Ca o recomandare generală, toate caracterele dintr-un paragraf trebuie să aibă aceeași mărime a fontului.

35. Paragrafele de text care conțin subscript sau superscript trebuie să folosească o mărime de font de cel puțin 12 (14 fiind recomandat - cu cât mai mare, cu atât mai bine). Cutia limitatoare a caracterelor la subscript sau superscript trebuie să fie suficient de aproape de cea a caracterelor normale de pe același rând (acest lucru previne ca procedurile de OCR-are să pună subscriptul/superscriptul pe rânduri diferite).

36. Fonturile recomandate sunt următoarele, în ordine:

- (a) Familia Monospaced: OCR-B, Courier New, Free Mono.
- (b) Familia Serif: ITC Officina Serif, Times New Roman, Free Times.
- (c) Familia Sans Serif: Verdana, ITC Officina Sans, Arial, Helvetica, DejaVu Sans.

Cu toate acestea, fonturile Arial, Helvetica, DejaVu Sans, Free Times, și Times New Roman nu sunt recomandate pentru cererile care conțin formule chimice și/sau matematice, cât și acronime care combină literele cu cifre. Pentru caractere chinezești, este recomandat fontul Song.

37. Caracterele fonturilor trebuie să fie bine formate, fără umbre. Spațiile dintre caractere trebuie să fie suficient de mari (spațierea îngustă trebuie evitată).

38. Fonturile înguste sau cursive nu trebuie folosite.

39. Stilurile bold și italic trebuie evitate cât de mult posibil.

40. Caracterele neobișnuite (non-standard/non-tipice/neregulate) trebuie evitate cât de mult posibil. Dacă sunt necesare, acestea trebuie să aparțină alfabetului grecesc standard sau fontului de simboluri (în ordinea preferată). Caracterele care nu pot fi găsite în aria UNICODE nu trebuie folosite; aceste caractere sunt recunoscute de programele de OCR-are ca imagini încorporate și, în consecință, fac textul recunoscut greu de citit. Fiecare oficiu își va formula și publica cerințele proprii legate de setul de caractere care pot fi utilizate pentru pregătirea cererii de brevet.

41. Textul nu trebuie să fie subliniat. Dacă este necesar, trebuie să se asigure faptul că sublinierea nu intersectează cutiile limitatoare ale caracterelor subliniate.

RECOMANDĂRI PENTRU COMBINAȚIILE DINTRE LIMBI

42. În cadrul secțiunilor/paginilor cererilor de invenție, amestecurile dintre limbile asiatice (de exemplu, bazate pe ideograme) și europene (de exemplu, alfabet latin și chirilic) creează probleme pentru procedurile de OCR-are, și trebuie evitate, cu excepția situațiilor în care sunt necesare.

RECOMANDĂRI DE SCANARE

43. Cererile de brevet trebuie să fie scanate ori în alb-negru ori în nuanțe de gri.

44. Rezoluția preferată pentru scanare este de 300 dpi. Scanarea la rezoluții mai mici de 300 dpi, chiar și în nuanțe de gri, poate da naștere unor documente de calitate slabă, publicate de oficii, deoarece schimbul de documente dintre oficii și procesele de publicare implică de obicei conversii în pagini TIFF de grup IV în alb-negru, la rezoluția de 300 dpi.

45. Documentele scanate trebuie convertite în format PDF ori TIFF.

CORECTURI

46. Corecturile textului unei cereri trebuie să fie făcute cu reimprimarea întregii pagini. Semne doveditoare de corectură - cum sunt, de exemplu, cele specificate în Standardul internațional ISO 5776 - nu sunt acceptate. Corectura cu pastă albă corectoare, bandă de hârtie auto-adezivă, ștergere sau tăiere nu este acceptată. Paginile de înlocuit nu trebuie trimise la oficiu prin fax folosind rezoluția standard de fax; paginile trebuie trimise ori fizic sau prin fax, folosind o rezoluție mai mare sau egală cu 300 dpi, ori prin orice alt mod de transfer în rețea care este acceptat de oficiu, cu condiția ca fiecare pagină să fi fost scanată la o rezoluție mai mare sau egală cu 300 dpi.

RECOMANDĂRI PENTRU OFICII

47. Oficiile de brevete trebuie să evite modificarea paginilor primite înainte de a le preda pentru scanare și pentru operațiile de OCR-are. De exemplu, unele practici actuale includ operații de ștampilare, care pot să suprapună caractere pe pagini, făcând textul predat de solicitant nelizibil pentru procedurile de OCR-are. Dacă ștampilele/modificările trebuie să fie aplicate pe paginile originale, oficiul trebuie să ia măsuri pentru a se asigura că modificările apar doar în marginile documentelor, cum este definit în Anexa 1.

48. Pe viitor, oficiile de brevete trebuie să evite elaborarea de formulare de hârtie pentru comunicarea dintre solicitanți și oficiu. Din experiențe anterioare s-a demonstrat că elaborarea și stabilirea unor sisteme de formulare on-line sunt preferate față de construirea unor sisteme care să recunoască formularele de hârtie. Cu toate acestea, se fac următoarele recomandări pentru elaborarea de formulare de hârtie, în scopul facilitării recunoașterii acestora:

- Liniile punctate nu trebuie folosite în formulare pentru a indica utilizatorului unde trebuie să introducă textul.
- Se pot alege culori (de exemplu, gri deschis) pentru marcarea box-urilor conținând caractere (de exemplu, ecuații).
- Culorile nu trebuie folosite în zonele care urmează să primească elemente gri sau colorate, cum ar fi o semnătură sau o figură.

IMPLEMENTARE

49. Se recomandă ca Oficiile care intenționează să înceapă să accepte sau să ceară arhivarea cererilor de brevet scrise în format OCR, să publice un ghid complet în Revistele lor oficiale, la intervale regulate, și pe paginile lor web, în care să definească tipul/tipurile exacte de caractere care sunt permise, și să specifice mărimea exactă a hârtiei care este admisă.

Exemple

50. Exemple de bune și rele practici privind prelucrarea OCR sunt reproduse în Anexa 2 a acestei Recomandări. Exemplele arată ceea ce trebuie și ceea ce nu trebuie făcut, împreună cu scurte explicații.

[Urmează Anexele]

ANEXA 1

[Textul aplicației trebuie să fie scris în acest chenar.
Numerele de pagină, numerele de rând și adnotările
trebuie inserate în margini.]

Dimensiunea originală = A4

ANEXA 2

EXEMPLE DE BUNE ȘI RELE PRACTICI

În această anexă veți găsi exemple bune și exemple rele de pagini de documente de brevet, referitoare la acuratețea obținută când s-au efectuat operații de OCR-are asupra lor.

Exemple de bune practici: Exemplul 1: o pagină bună de descriere

WO 2006/111319

PCT/EP2006/003401

Projection exposure system, method for manufacturing a micro-structured structural member by the aid of such a projection exposure system and polarization-optical element adapted for use in such a system

5

The invention relates to a projection exposure system, in particular for micro-lithography. The invention further relates to a method for manufacturing a micro-structured component and a polarization-optical element for the extreme ultraviolet (EUV) region.

10

For highest possible precision of the optical image to be obtained in complicated optical instruments such as a projection exposure system, the influence of the polarization of the light must be considered or, respectively, the polarization must be influenced specifically. For example, in particular in case of great incidence angles, polarization effects occur in the mirror systems, which projection exposure systems in the EUV region are based on, for lack of suitable transparent materials. These polarization effects are in particular due to the varying reflectivity of the mirrors for s-polarized and p-polarized light and can give rise to imaging errors or other undesired effects. Efforts have been made to measure possible polarization effects in the individual components of projection exposure systems.

15

For example, EP 1 306 665 A2 discloses an optical instrument for measuring polarization-dependent properties which comprises a light source in the EUV or X-radiation region and a rotatable polarizer. The polarizer is substantially comprised of a set of mirrors that reflects the incident light at least three times. The mirrors are arranged in such a way that the optical axes of the incident and emergent light are on the same straight line.

20

25

Exemplul 2: o pagină bună de revendicări

WO 2008/015644

PCT/IB2007/053030

5

CLAIMS

1. A sports or games apparatus (10), characterized in that it comprises a flexible grid (11) of rectangular shape, which has two base sides (12) and two height sides (13), and is formed from a plurality of grid elements (20, 30, 40), in which the said grid elements comprise a pair of rigid bars (20) forming respectively the said two height sides of the grid, capable of being attached to corresponding support structures (S); a pair of elastic base cords or tapes (30), forming respectively the said two base sides of the grid, with their ends (31) attached to the rigid bars; and a plurality of elastic grid cords or tapes (40), the ends of each of the said elastic grid cords or tapes being attached to another two corresponding grid elements, in such a way that, in an installed condition of the said apparatus, in which the said rigid bars are attached to the said support structures in such a way as to bring the said elastic base cords or tapes and the said elastic grid cords or tapes into tension, the said grid elements are positioned so as to form a grid with a rectangular mesh.
2. An apparatus according to Claim 1, in which the ends (31, 41) of each elastic base cord or tape (30) and of each elastic grid cord or tape (40) are movable along the direction of extension of the corresponding grid element to which they are attached.
3. An apparatus according to Claim 2, in which the ends (31, 41) of each elastic base cord or tape (30) and of each elastic grid cord or tape (40) are bent back to form a noose and attached to themselves, and the corresponding grid elements to which they are attached in a movable way are inserted into the corresponding nooses (42) formed by them.
4. An apparatus according to any one of the preceding claims, in which at least one elastic grid cord or tape (40) comprises at least one intermediate point (43) attached to another elastic grid cord or tape (40).
5. An apparatus according to Claim 4, in which at least one intermediate point is movable along the direction of extension of the corresponding elastic grid cord or tape (40) to which it is attached.

Toate recomandările sunt luate în calcul: margini, un font standard (Times New Roman), o mărime bună de font, niciun număr de rând, utilizare limitată de bold, fără italic, fără text subliniat; numerotarea revendicărilor este adecvată și bine separată de textul revendicărilor.

Anexa 2, pagina 3

Exemplul 3: o pagină complexă bună de descriere

WO 2006/102655

PCT/US2006/011076

[0134] When performing the first iteration of step S9-4, the values of D_{25} , A_{25} , D_b and A_b are the values previously calculated at step S7-2, while all values of λ_n are zero.

[0135] The equations used by solver 244 at step S9-6 comprise the following in this embodiment:

$$\text{if } (\lambda_{x,y,z-\max}^{n+1})_{ang \& lin} < 0 \text{ then } \lambda_{x,y,z-\max}^{n+1} = 0 \quad (46)$$

$$\text{if } (\lambda_{x,y,z-\min}^{n+1})_{ang \& lin} > 0 \text{ then } \lambda_{x,y,z-\min}^{n+1} = 0 \quad (47)$$

$$\lambda_{lin}^{n+1} = \lambda_{lin-\min}^{n+1} + \lambda_{lin-\max}^{n+1} \quad (48)$$

$$\lambda_{ang}^{n+1} = \lambda_{ang-\min}^{n+1} + \lambda_{ang-\max}^{n+1} \quad (49)$$

[0136] The equations used by solver 244 at step S9-8 comprise the following in this embodiment:

$$D_i^{n+1} = D_i^n + L \frac{(\lambda_{lin}^{n+1} - \lambda_{lin}^n)}{m_a} \quad (50)$$

$$A_a^{n+1} = A_a^n + I_a^{-1} [r_a^s] L (\lambda_{lin}^{n+1} - \lambda_{lin}^n) + I_a^{-1} T (\lambda_{ang}^{n+1} - \lambda_{ang}^n) \quad (51)$$

$$D_b^{n+1} = D_b^n - L \frac{(\lambda_{lin}^{n+1} - \lambda_{lin}^n)}{m_b} \quad (52)$$

$$A_b^{n+1} = A_b^n - I_b^{-1} [r_b^s] L (\lambda_{lin}^{n+1} - \lambda_{lin}^n) - I_b^{-1} T (\lambda_{ang}^{n+1} - \lambda_{ang}^n) \quad (53)$$

[0137] Referring again to Figure 7, at step S7-6, solver 244 performs a convergence test. In this embodiment, solver 244 performs processing to determine whether the values of λ calculated for the current iteration differ from the values of λ calculated for the previous iteration by more than a predetermined threshold, in accordance with the following equation:

$$\sum_{\lambda} \frac{(\lambda^{n+1} - \lambda^n)^2}{\lambda^{n2}} \leq \text{Threshold} \quad (54)$$

[0138] In this embodiment, the threshold employed in Equation (54) is set to 10^{-4} .

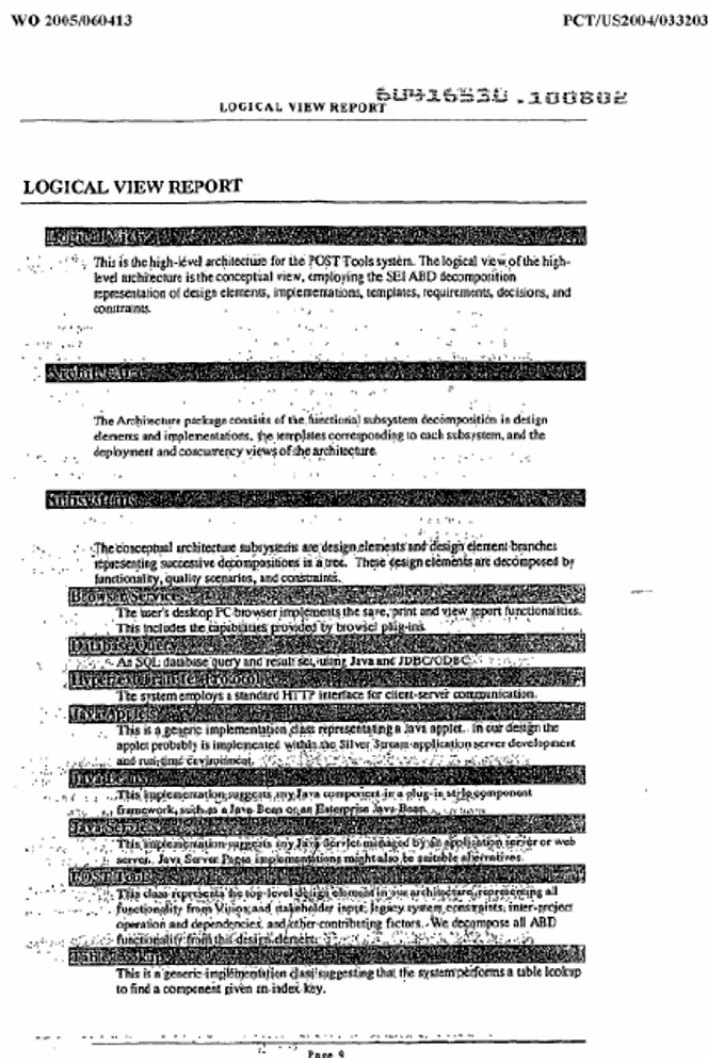
[0139] At step S7-8, solver 244 determines whether a predetermined number of iterations of the processing at steps S7-2 to S7-8 have been performed. In this embodiment, solver 244 determines whether 50 iterations have been performed.

Paragrafele de text nu conțin caractere matematice neobișnuite. Formulele matematice sunt introduse corect, lăsând posibilitatea unei segmentări ușoare a imaginilor încorporate, de către programul de OCR-are. Un rezultat posibil al segmentării este arătat pe albastru.

Anexa 2, pagina 4

Exemple de rele practici

Exemplul 1: o pagină de proastă calitate, cu multe deficiențe



78

Acest exemplu nu este conform cu paragraful 10 (pagina a fost probabil predată oficiului prin fax, la 200 dpi - vezi petele - și părți din text au fundal gri închis). Exemplul nu este conform nici cu paragrafele 13 și 47: un număr de referință (604115530.100802) este ștampilat în partea de mijloc a paginii (ar trebui să fie pe margine). Numerotarea paginilor este incorectă (ar trebui să fie "9", nu "page 9", vezi paragraful 34). În ultimul rând, mărimea de font este prea mică (paragraful 15). Astfel de pagini ar trebui, în mod ideal, să nu fie acceptate de oficii, și să se ceară pagini care să le înlocuiască (pagina este imposibil de OCR-at corect).

Anexa 2, pagina 5

Exemplul 2: o pagină cu un fundal care nu este alb

WO 2005/097403

- 13 -

PCT/FR2005/050194

REVENDICATIONS

1. Dispositif d'usinage (D) du type de celui associant une machine-outil d'usinage (100) à un dispositif porte-pièce (200) équipé d'un axe de mise en mouvement de rotation transversal (A) par rapport à l'axe de plongée (Z), CARACTÉRISÉ PAR LE FAIT QUE le dispositif porte-pièce (200) est constitué par un bâti (210) supportant deux paliers de guidage (210 et 230) en rotation selon ledit axe de rotation transversal (A), la structure formée par le bâti (210) et les deux paliers (220 et 230) étant fermée par la pièce à usiner (300) dont les extrémités viennent se fixer auxdits paliers (310 et 220), la pièce à usiner (300) étant une pièce longue du type de celle comportant des surfaces à usiner concentrées à ses deux extrémités ET PAR LE FAIT QUE la machine-outil (100) est du type de celle assurant la mise en mouvement de deux coulants porte-outil indépendants (110 et 120) de façon à ce que les usinages des deux extrémités de la pièce (300) soient réalisés par un coulant différent.

2. Dispositif d'usinage (D) selon la revendication 1, CARACTÉRISÉ PAR LE FAIT QUE chaque palier (220 et 230) comprend et guide un plateau tournant (221 et 231) équipé d'un moyen de mise en mouvement motorisé, la rotation des deux plateaux (221 et 231) étant synchronisée.

3. Dispositif (D) selon la revendication 2, CARACTÉRISÉ PAR LE FAIT QUE chaque plateau (220 et 230) est équipé de deux appuis (410, 420 et 510, 520) pour accueillir et maintenir en position l'extrémité de la pièce (300).

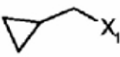
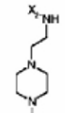
4. Dispositif (D) selon la revendication 1, CARACTÉRISÉ PAR LE FAIT QUE le bâti (210) du dispositif porte-pièce (200) est lui-même monté mobile en rotation selon un axe (B) perpendiculaire à l'axe (A) de rotation défini par les deux paliers (220 et 230) qu'il supporte.

5. Procédé d'usinage d'une pièce longue (300) du type de celle comportant des surfaces à usiner concentrées à

Acest exemplu nu este conform cu paragraful 12. Pagina trebuie trecută printr-un filtru pentru a elimina fundalul cu pete, înainte de a fi predată pentru o operația de OCR-are. Dacă se OCR-ează așa cum este, textul obținut nu va fi lizibil.

Anexa 2, pagina 6

Exemplul 3: o pagină având caractere pale

#	R2	A	UV max [nm]:	MS (ESI) (M+H) ⁺	
25			305, 350	476	Trihy 1,41 ((m, 2H (m, 1H

Beispiele 26-40

Die folgenden Verbindungen sind über ein analoges
beschrieben, hergestellt. Die Herstellung des Benz
5 beschrieben. Das für die Darstellung des Amids ei

O porțiune mică a paginii este mărită pentru a evidenția caracterele: culoarea textului original este probabil gri, dând naștere, după o scanare în 300 dpi alb-negru, unor caractere care nu sunt solide. În consecință, acuratețea textului OCR-at este slabă (acest exemplu nu este conform cu paragraful 12).

Exemplul 4: o pagină cu text scris de mână

TITLED : JIG HEAD SWAY BAR

BACK GROUND

IN THE ART OF FISHING THERE IS A PIECE
OF TACKLE KNOWN AS A PIVOT - HEAD JIG WHICH
USES SPECIALIZED OR SPECIFICALLY SHAPED HOOKS TO
PROVIDE AN ACTION PRODUCING LURE COMBINATION.
MY INVENTION THE SWAYBAR ALLEVIATES THIS
NEED FOR SPECIAL HOOKS BY BEING ABLE
TO BOTH SUPPORT THE JIG HEAD AND ALLOW
FOR CONNECTION OF OTHER REQUIRED TACKLE

După cum era de așteptat, textul obținut în urma OCR-ării acestei pagini nu este lizibil. Oficiile trebuie să ceară texte scrise la mașină, pentru a se asigura calitatea minimă de publicare.

Anexa 2, pagina 7

Exemplul 5: o pagină cu o configurație nerecomandată și alte deficiențe

WO 2005/086760

PCT/US2005/007335

38

relation to the determination of AN by FTIR spectroscopy

This concept is illustrated in Figure 1 for AN, the BN analysis being analogous but using a different reagent. Differential spectroscopy is then used to eliminate the spectral contributions from the base oil and any additives and/or contaminants and breakdown products present in the oil that may spectrally interfere with the measurement of the signal from the reaction product. This is achieved by treating a portion of the sample with a blank reagent, this portion effectively serving as a reference oil. Figure 2 illustrates the general analytical protocol.

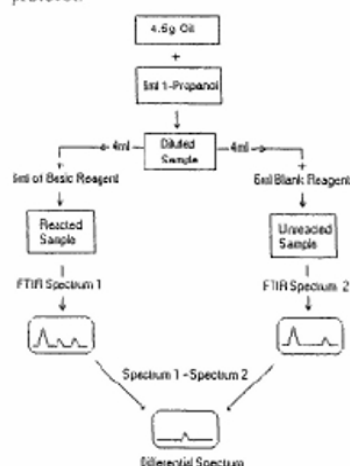


Figure 2. Analytical protocol for the determination of AN by FTIR spectroscopy.

In this procedure, the sample is first diluted with an innocuous solvent (1-propanol), then split and treated with a

reactive and a blank reagent to produce two samples for spectral analysis. Since these two samples are the same except for the reaction products, subtraction of their spectra leaves only the spectral contribution related to AN.

The COAT AN/BN Analyzer

The COAT AN/BN Analyzer has been designed and programmed to automate AN/BN analyses based on the concepts laid out above. Figure 3 illustrates key components of the COAT AN/BN Analyzer: an FTIR spectrometer, a sample handling accessory, an autosampler, and the computer that controls the system.

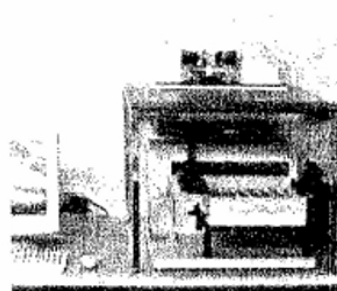


Figure 3. The COAT AN/BN Analyzer and its key components

The compact nature of the sample handling system is made possible by the dilution of the sample in the analytical protocol (Figure 2), allowing a micropump to be substituted for the peristaltic pump employed in most FTIR used oil analyzers. The resulting low viscosity of the sample dramatically

3

Această pagină nu este conform cu următoarele recomandări: paragraful 17 (format cu o singură coloană), paragraful 39 (sunt folosite fonturile italic și bold), paragraful 46 (s-au făcut corecturi manuale după imprimare). Alinierea la dreapta și la stânga a paragrafului este, de asemenea, nerecomandată (paragraful 30), deși, în acest caz, aceasta nu va avea efecte negative la OCR-are, deoarece cuvintele încă sunt separate suficient de spații albe. În ultimul rând, acest exemplu nu este conform nici cu paragraful 27 (spațiere la un rând și jumătate).

Anexa 2, pagina 8

Exemplul 6: o pagină cu numere de rând care sunt prea mici

	WO 2004/110497	PCT/US2004/013820
	[0028] Figs. 9A-9B are plots showing the percent of mitomycin C released from liposomes comprised of HSPC/mPEG-DSPE/lipid-DTB-mitomycin C (Fig. 9A) and HSPC/cholesterol/mPEG-DSPE/lipid-DTB-mitomycin C (Fig. 9B) as a function of time of incubation in the presence of cysteine at concentrations of 150 μM (closed symbols) and at 1.5 mM (open symbols);	
5	[0029] Fig. 10 is a plot of growth rate of M109 cells, expressed as a percentage based on growth of M109 cells in the absence of drug and cysteine, as a function of mitomycin C amount, in nM, for free mitomycin C (open triangles), liposomes comprised of HSPC/mPEG-DSPE/lipid-DTB-mitomycin C (closed squares), and liposomes comprised of HSPC/cholesterol/mPEG-DSPE/lipid-DTB-mitomycin C (open circles);	
10	[0030] Fig. 11A is a plot of growth rate of M109 cells, expressed as a percentage based on growth of M109 cells in the absence of drug or cysteine, as a function of mitomycin C concentration in nM. Shown are cells treated mitomycin C in free form (open triangles) and with mitomycin C in free form plus 1000 μM cysteine (closed triangles). Also shown are cells treated with the liposome formulation comprised of HSPC/PEG-DSPE/lipid-DTB-mitomycin C (open circles) and with the liposome formulation with additional cysteine added at concentrations of 150 μM (open diamonds), 500 μM (closed circles) and 1000 μM (open squares);	
15	[0031] Fig. 11B is a plot of growth rate of M109 cells, expressed as a percentage based on growth of M109 cells in the absence of drug or cysteine, as a function of mitomycin C concentration in nM. Shown are cells treated mitomycin C in free form (open triangles) and with mitomycin C in free form plus 1000 μM cysteine (closed triangles). Also shown are cells treated with the liposome formulation comprised of HSPC/cholesterol/mPEG-DSPE/lipid-DTB-mitomycin C (open circles) and with the liposome formulation with additional cysteine added at concentrations of 150 μM (open diamonds), 500 μM (closed circles) and 1000 μM (open squares);	
20	[0032] Fig. 12 is a plot showing the percent increase in cytotoxicity (as determined by $(IC_{50, no\ cysteine}/IC_{50, cysteine}) \times 100$) of free mitomycin C (closed squares), mitomycin C associated with liposomes comprised of HSPC/cholesterol/mPEG-DSPE/lipid-DTB-mitomycin C (closed circles), and liposomes comprised of HSPC/mPEG-DSPE/lipid-DTB-mitomycin C (open triangles) to M109 cells <i>in vitro</i> at various concentrations of cysteine;	
25	[0033] Fig. 13A is a plot showing the concentration of mitomycin C in the blood of	
30		

6

Numerele de rând sunt o problemă pentru programele de OCR-are din mai multe motive (paragraful 14):

- pot să nu fie aliniate cu rândurile cărora le corespund, ceea ce duce la detecția eronată a rândurilor de bază;
- pot fi prea mici, ceea ce duce la erori de recunoaștere care vor împiedica procedurile de extragere XML de la eliminarea lor completă din corpul de text al paginii;
- pot fi așezate greșit în partea de text a paginii, sau pe margini, dar prea aproape de corpul textului și, în consecință, vor apărea înăuntru fluxului de text exportat de operațiile de OCR-are.

În acest exemplu, numerele sunt prea mici.

Caracterele subscript sunt, de asemenea, prea mici în acest exemplu (paragraful 35).

Anexa 2, pagina 9

Exemplul 7: o pagină care conține text scris în mai multe direcții.

WO 2005/081642

PCT/JP2005/003688

Table 11 (continued-2)

		Amount in retardation-controlling agent solution (mass parts)					Amount in UV absorber solution (mass parts)						Mixing ratio of solutions			
		Retardation-control agent A-2	Retardation-control agent A-12	Retardation-control agent E	Retardation-control agent C	Retardation-control agent D	UV absorber A	UV absorber B	UV absorber C	UV absorber D	UV absorber E	UV absorber F	Cellulose acetate solution	Mating agent solution	Retardation-controlling agent solution	UV absorber solution
This invention	Cellulose acetate film 25			6	14							15	94.6	1.2	6.2	1.3
This invention	Cellulose acetate film 26	3		3	14							15	94.6	1.2	7.0	3.2
This invention	Cellulose acetate film 27	5		5	10							15	94.6	1.2	6.2	0.8
This invention	Cellulose acetate film 28			5		15	4.8	10.2					94.6	1.2	6.2	0.8
This invention	Cellulose acetate film 29		10			10	4.8	10.2					94.6	1.2	6.2	0.8
This invention	Cellulose acetate film 30					15	4.8	10.2					94.6	1.2	6.2	0.8
Comparative example	Cellulose acetate film 31	10	10										94.6	1.2	6.6	0
Comparative example	Cellulose acetate film 32			20									94.6	1.2	4.1	0
Comparative example	Cellulose acetate film 33									5	10		94.6	1.2	0	6.3
Comparative example	Cellulose acetate film 34	10	10				10.5	4.5					94.6	1.2	7.1	0.8
Comparative example	Cellulose acetate film 35	10	10				10.5	4.5					94.6	1.2	7.1	0.8

133

Acest exemplu nu este conform cu paragraful 20.

Una dintre limitările celor mai bune motoare de OCR-are disponibile în momentul de față este aceea că ele nu pot să citească decât un text scris într-o singură direcție pe o pagină (se face o procesare anterioară a paginii, pentru a detecta orientarea principală a paginii). În consecință, toate cuvintele care nu sunt scrise pe direcția principală a textului sunt ignorate. Se acceptă, bineînțeles, să apară pe o pagină un tabel având orientarea tip vedere, sau chiar un text principal orientat tip vedere, cu adnotări pe margini în orientarea tip portret (numărul paginii, numărul cererii etc.).

Exemplul 8: o pagină cu formule matematice amestecate cu textul

WO 2005/116630 PCT/US2005/017216

19

$$\Delta \mathbf{L} = \frac{\hbar}{2} - \mathbf{r} \times e \mathbf{A} \quad (33)$$

$$= \left[\frac{\hbar}{2} - \frac{e \phi}{2\pi} \right] \hat{z} \quad (34)$$

In order that the change of angular momentum, $\Delta \mathbf{L}$, equals zero, ϕ must be $\Phi_0 = \frac{h}{2e}$, the magnetic flux quantum. The magnetic moment of the electron is parallel or

5 antiparallel to the applied field only. During the spin-flip transition, power must be conserved. Power flow is governed by the Poynting power theorem,

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\frac{\partial}{\partial t} \left[\frac{1}{2} \mu_0 \mathbf{H} \cdot \mathbf{H} \right] - \frac{\partial}{\partial t} \left[\frac{1}{2} \epsilon_0 \mathbf{E} \cdot \mathbf{E} \right] - \mathbf{J} \cdot \mathbf{E} \quad (35)$$

Eq. (36) gives the total energy of the flip transition which is the sum of the energy of reorientation of the magnetic moment (1st term), the magnetic energy (2nd term), the

10 electric energy (3rd term), and the dissipated energy of a fluxon treading the orbitsphere (4th term), respectively,

$$\Delta E_{\text{mag}}^{\text{spin}} = 2 \left(1 + \frac{\alpha}{2\pi} + \frac{2}{3} \alpha^2 \left(\frac{\alpha}{2\pi} \right) - \frac{4}{3} \left(\frac{\alpha}{2\pi} \right)^2 \right) \mu_B B \quad (36)$$

$$\Delta E_{\text{mag}}^{\text{spin}} = g \mu_B B \quad (37)$$

15 where the stored magnetic energy corresponding to the $\frac{\partial}{\partial t} \left[\frac{1}{2} \mu_0 \mathbf{H} \cdot \mathbf{H} \right]$ term increases, the stored electric energy corresponding to the $\frac{\partial}{\partial t} \left[\frac{1}{2} \epsilon_0 \mathbf{E} \cdot \mathbf{E} \right]$ term increases, and the $\mathbf{J} \cdot \mathbf{E}$ term is dissipative. The spin-flip transition can be considered as involving a magnetic moment of g times that of a Bohr magneton. The g factor is redesignated the fluxon g factor as opposed to the anomalous g factor. Using $\alpha^{-1} = 137.03603(82)$, the

20 calculated value of $\frac{g}{2}$ is 1.001 159 652 137. The experimental value [23] of $\frac{g}{2}$ is 1.001 159 652 188(4).

1.G. SPIN AND ORBITAL PARAMETERS

The total function that describes the spinning motion of each electron orbitsphere

25 is composed of two functions. One function, the spin function, is spatially uniform over the orbitsphere, spins with a quantized angular velocity, and gives rise to spin angular momentum. The other function, the modulation function, can be spatially uniform—in which case there is no orbital angular momentum and the magnetic moment of the electron orbitsphere is one Bohr magneton—or not spatially uniform—in which case

30 there is orbital angular momentum. The modulation function also rotates with a quantized angular velocity.

The spin function of the electron corresponds to the nonradiative $n=1, \ell=0$

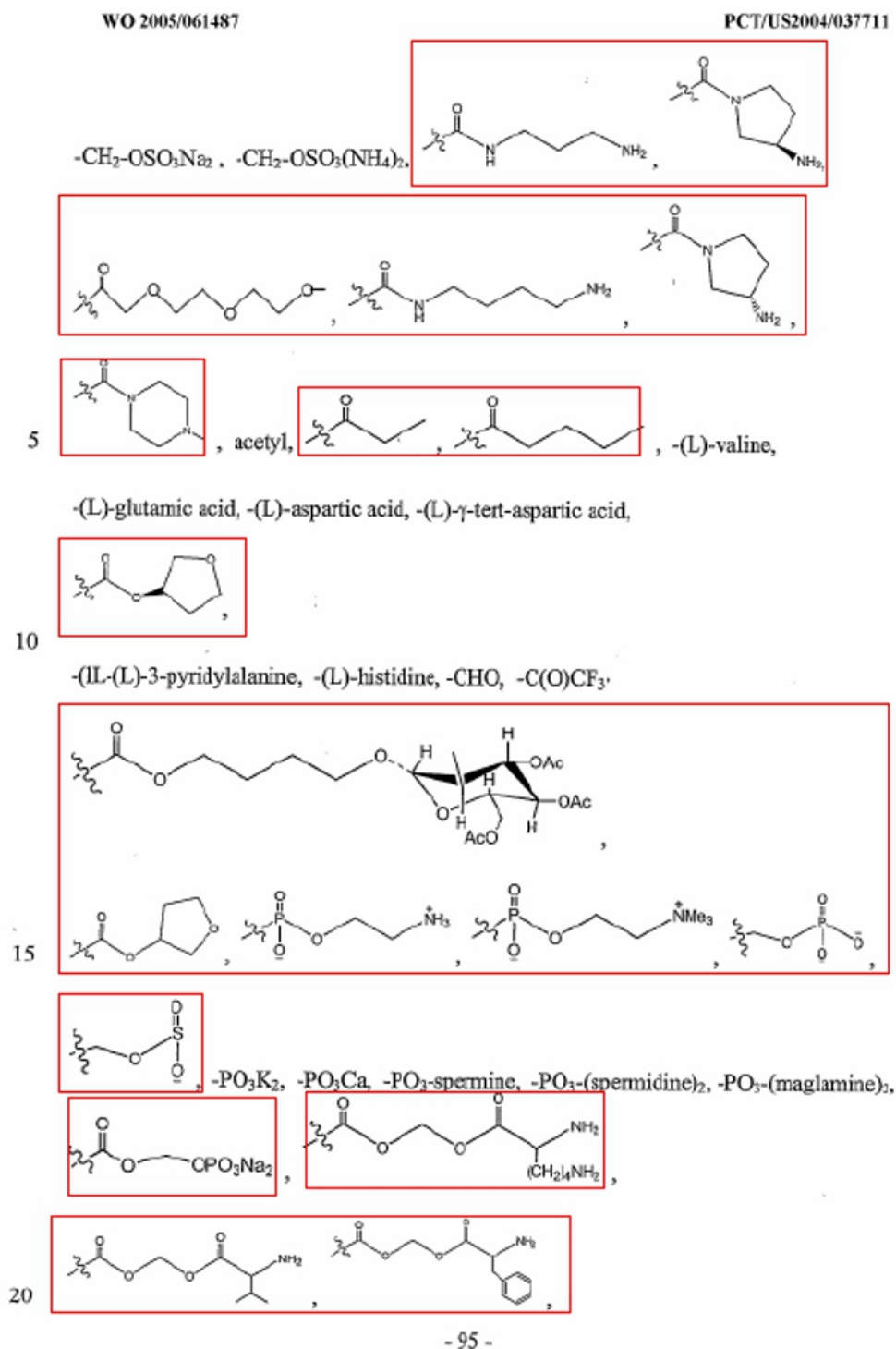
Acest exemplu nu este conform cu paragraful 23. Motorul de OCR are nu este capabil să separe corect textul de formule (vezi rezultatul unei segmentări manuale a formulelor, pe roșu: formulele încorporate se și intersectează).

Ca observație generală, în acest exemplu textul și formulele sunt prea compacte pentru a fi recunoscute bine; acest exemplu nu corespunde nici cu paragrafele 27 și 28.

Acest exemplu folosește, de asemenea, caractere neobișnuite: simbolurile grecești pot fi folosite chiar dacă ele măresc gradul de dificultate al recunoașterii paginii (vezi paragraful 40). Cu toate acestea, se recomandă insistent ca fonturile italice, bold sau subliniate să nu fie combinate cu caracterele neobișnuite (paragraful 39).

Anexa 2, pagina 11

Exemplul 9: o pagină cu formule chimice încorporate, amestecate cu textul



Acest exemplu nu este conform cu paragraful 23. Pe roșu, se poate găsi un rezultat presupus al segmentării figurilor (făcută manual). Această segmentare nu poate fi făcută corect de un program de OCR-are pentru că formulele sunt prea apropiate de textul care le înconjoară.

Exemplul 10: o pagină cu caractere subscript care sunt prea mici

WO 2005/110416

PCT/US2005/015897

R_1 is hydrogen, C_1 - C_6 alkyl, C_2 - C_6 alkenyl, C_1 - C_6 alkynyl, C_1 - C_6 alkoxy, C_1 - C_6 haloalkyl, C_1 - C_6 haloalkoxy, $(C_3$ - C_7 cycloalkyl) C_3 - C_6 alkyl;

R_2 is selected from alkoxy, cycloalkoxy, phenyl, 4- to 7-membered heterocycles, $-O(CH_2)_n$ phenyl, $-O(CH_2)_n$ pyridyl, $-E-(CR_3R_4)_n-Q$, and Q , each of which is substituted with between 0 and 3 substituents selected from halogen, cyano, hydroxy, oxo, $(CR_3R_4)_2-T$, C_1 - C_6 alkyl, C_1 - C_6 alkoxy, C_1 - C_6 haloalkyl, C_1 - C_6 haloalkoxy, mono- and di- $(C_1$ - C_6 alkyl)amino, $(C_1$ - C_6 alkyl) $(CR_3R_4)_2-T$ amino, benzyl, $S(O)_j(C_1$ - C_6 alkyl), α,ω - C_1 - C_6 alkylene, α,ω - C_1 - C_6 alkyleneoxy, α,ω - C_1 - C_6 alkylenedioxy, $-E-(CH_2)_n-Q$, and Q ;

T is CO_2H , $CONH_2$, C_1 - C_6 alkoxycarbonyl, mono- or di- $(C_1$ - C_6 alkyl)aminocarbonyl, SO_3H , SO_3NH_2 or $SO_2(C_1$ - C_6 alkyl);

j is an integer ranging from 0 to 6;

Q is a saturated heterocyclic ring comprising between 4 and 7 ring members, in which the point of attachment is a carbon or nitrogen atom;

E is O , NR_5 , or a single covalent bond;

R_4 and R_5 are independently chosen from hydrogen, halogen, hydroxy, C_1 - C_6 alkyl, C_1 - C_6 alkenyl, $(C_3$ - C_7 cycloalkyl) C_3 - C_6 alkyl and C_1 - C_6 alkoxy; and

Ar is phenyl which is mono-, di-, or tri-substituted; or 1-naphthyl, 2-naphthyl, pyridyl, pyrimidinyl, pyrazinyl, pyridizyl, thienyl, thiazolyl, pyrazolyl, imidazolyl, tetrazolyl, oxazolyl, isoxazolyl, pyrrolyl, furanyl, indolyl, indazolyl, or triazolyl, each of which is optionally mono-, di-, or tri-substituted.

Yet other compounds of Formula VIII include those compounds in which the group designated:



is chosen from naphthyl, tetrahydronaphthyl, benzofuranyl, benzo[dioxolyl], indanyl, indolyl, indazolyl, benzo[dioxolyl], benzo[1,4]dioxanyl and benzoxazolyl, each of which is substituted with from 0 to 3 substituents independently chosen from R_4 .

Certain compounds of Formula IX include those in which

Ar is mono-, di-, or tri-substituted phenyl, which phenyl group is substituted with one to three substituents independently chosen from hydroxy, halogen, cyano, amino, nitro, $-COOH$, aminocarbonyl, $-SO_2NH_2$, C_1 - C_6 alkyl, C_1 - C_6 alkenyl, C_1 - C_6 alkynyl, C_1 - C_6 haloalkyl, C_1 - C_6 aminoalkyl, C_1 - C_6 hydroxyalkyl, C_1 - C_6 carboxyalkyl, C_1 - C_6 alkoxy, C_1 - C_6 haloalkoxy, C_1 - C_6 alkylthio, C_1 - C_6 alkanoyl, C_1 - C_6 alkanoyloxy, C_3 - C_7 alkanone, C_1 - C_6 alkyl ether, mono- or di- $(C_1$ - C_6 alkyl)amino C_3 - C_6 alkyl, $-NHC(=O)(C_1$ - C_6 alkyl), $-N(C_1$ - C_6 alkyl) $C(=O)(C_1$ - C_6 alkyl), $-NHS(O)_j(C_1$ - C_6 alkyl), $-(C_1$ - C_6 alkyl) $C(=O)NH_2$, $-(C_1$ - C_6 alkyl) $C(=O)NH(C_1$ - C_6 alkyl), $-(C_1$ - C_6 alkyl) $C(=O)NH(C_1$ - C_6 alkyl)(C_1 - C_6 alkyl), $-S(O)_j(C_1$ - C_6 alkyl), $-S(O)_jNH(C_1$ - C_6 alkyl), $-S(O)_jN(C_1$ - C_6 alkyl)(C_1 - C_6 alkyl) and Z ; or

Acesta este un exemplu tipic în care caracterele subscript sunt prea mici pentru a se asigura o recunoaștere corectă. Acest fenomen este întâlnit frecvent în brevetele din domeniul chimiei.

Anexa 2, pagina 13

Exemplul 11: o pagină cu tabele într-un format greșit

WO 2005/063765

PCT/US2004/043492

Table D

Other compounds of the invention result from selecting appropriate features from the table of possible features below. For example, compound A77 results from the following selections: none-morpholino-aryl-OCH₂(CO)-piperazine-CH₃.

5	Left-hand substituent	Left-hand ring	Aryl or heteroaryl	Ring substituent	Nitrogen feature	Right-hand substituent
	CH ₃	morpholino	aryl	OCH ₂	NHM	alkyl
	isopropyl	piperazine	thiophene	OCH ₂ (CO)	NMM	alkoxy
	CH ₃ CH ₂ O(CO)CH ₂			SO ₂	morpholino	alcohol
	none			OCH ₂ (CO)OCH ₂	piperazine	substituted amine
					piperidine	acid
					pyrazole	ester
					pyrrolidine	CH ₂ CH ₂ OCH ₃
						CH ₂ CH ₂ OH
						CH ₂ NH ₂
						CH ₂ NHCH ₂ CH ₂ CH ₃
						CH ₂ NHCH ₃
						CH ₂ NHCHCH ₃ CH ₃
						CH ₃
						CHCH ₃ CH ₃
						COOCH ₂ CH ₃
						none

Table E

Other compounds of the invention result from selecting appropriate features from the table of possible features below. For example, compound B3 results from

10 the following selections: none-morpholino-aryl-CH₂-piperazine-CH₂CH₂OH.

	Left-hand substituent	Left-hand ring	Aryl or heteroaryl	Ring substituent	Nitrogen feature	Right-hand substituent
	CH ₃	morpholino	aryl	CH ₂	NHM	alkyl
	isopropyl	piperazine	thiophene	CH ₂ CH ₂	NMM	alkoxy
	CH ₃ CH ₂ O(CO)CH ₂			CH ₂ CH ₂ CH ₂	morpholino	alcohol
	none			CH ₂ CH ₂ CH ₂ CH ₂	piperazine	substituted amine
					piperidine	acid
					pyrazole	ester
					pyrrolidine	CH ₂ CH ₂ OCH ₃
						CH ₂ CH ₂ OH
						CH ₂ NH ₂
						CH ₂ NHCH ₂ CH ₂ CH ₃
						CH ₂ NHCH ₃
						CH ₂ NHCHCH ₃ CH ₃
						CH ₃

În acest exemplu, marginile tabelului lipsesc (nu este conform cu paragraful 33). În consecință, motorul de OCR-are va încerca să prelucreze conținutul tabelului ca un paragraf de text. Aceasta creează mai multe probleme:

- Mărimea fontului caracterelor din tabel este prea mică (paragrafele 34 și 35).
- Liniile de bază ale titlurilor de coloane sunt amestecate (paragraful 29). În consecință, programul va detecta greșit subscriptul și superscriptul.
- Textul obținut nu va lua în considerare coloanele:

Left-hand Left-Hand Aryl or Nitrogen

Substituent ring heteroaryl Ring substituent feature Righthand substituent CH₃...

Exemplul 12: o pagină aliniată

WO 2005/087962

1

PCT/EP2005/002168

GKSS-Forschungszentrum Geesthacht GmbH, Max-Planck-Stra-
Be 1, 21502 Geesthacht

Verfahren zur Herstellung von Profilen aus Leichtmetall-
werkstoff mittels Strangpressen

Beschreibung

Die Erfindung betrifft ein Verfahren zur Herstellung von
Profilen aus Leichtmetallwerkstoff, insbesondere Magne-
siumwerkstoff, mittels Strangpressen, bei dem ein Werks-
stoffvolumen durch eine Matrize, die die Form des ge-
wünschten Profils bestimmt, zur Ausbildung des Profils
gepreßt wird.

Die Herstellung von Profilen aus Leichtmetall- bzw.
Leichtmetall-Legierungswerkstoffen mittels eines Strang-
preßverfahrens ist eine allgemein eingeführte, bekannte
Technologie und wird industriell angewendet. So ist es
bekannt, daß konventionell verfügbare Leichtmetall- bzw.
Leichtmetall-Knetlegierungen in Form von Gußblöcken
durch konventionelles Strangpressen in Profilformen ge-
preßt werden. Dabei wird der Leichtmetall- bzw. Leicht-
metall-Legierungsblock, im folgenden zusammenfassend
kurz mit Werkstoffvolumen bezeichnet, bei Temperaturen

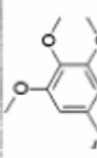
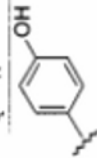
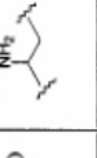
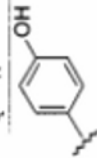
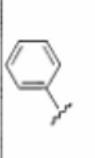
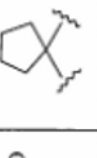
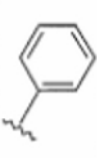
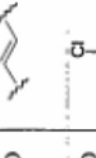
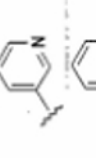
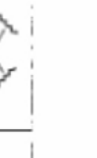
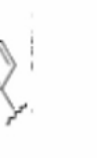
În acest exemplu, paragrafele sunt aliniate la stânga și la dreapta. Chiar dacă acest lucru face textul să pară mai estetic, uneori face și operațiile de OCR-are mai dificile, când pauzele dintre cuvinte sunt prea mici (paragraful 30). Acest exemplu nu este în conformitate nici cu paragraful 31, care spune că scindarea cuvintelor la capăt de rând trebuie evitată cât de mult posibil (pentru motorul de OCR-are este uneori dificil să facă distincția dintre cratime utilizate la cuvintele compuse și cele care despart cuvintele în silabe, în produsul final rezultând cuvinte care conțin cratime nedorite).

Exemplul 13: un tabel cu margini de calitate proastă

WO 2004/110415

- 60 -

PCT/EP2004/051048

Comp. No.	Exp. No.	Alk ^a	Y	Alk ^b	L	Physical data
106	B2	cb	C=O	-CH ₂ -		2R-trans
107	B3b	cb	C=O	-CH ₂ -		2R-trans
13	B8	cb	C=O			2R-trans, HCl(1:3); H ₂ O(1:1)
108	B2	cb	C=O			2R-trans HCl(1:2) H ₂ O(1:1)
109	B2	cb	C=O			2R-trans
110	B3b	cb	C=O			[2R-[2α,4β(E)]]
111	B2	cb	C=O			2R-trans

În acest exemplu, marginile tabelului din originalul primit anterior scanării sunt de calitate proastă. După scanare, procedeul de OCR-are nu poate să detecteze corect tabelul, și necesită o operație manuală de segmentare a paginii. Dacă o astfel de pagină nu este verificată de operator în ceea ce privește calitatea, textul final va conține caractere nedorite, care vor face indexarea documentului cu motoarele de căutare mai puțin eficace.

Anexa 2, pagina 16

Exemplul 14: caractere slabe la subscript și superscript

	WO 2005/100305	PCT/IB2005/006872
	-9-	
	thiazolyl, pyrazolyl, pyridinyl, pyrimidinyl, purinyl, quinolinyl, benzofuran and isoquinolinyl.	
5	p. "heteroaryl, optionally substituted," refers to a heteroaryl moiety as defined immediately above, in which up to 4 carbon atoms of the heteroaryl moiety may be substituted with a substituent, each substituent is independently selected from the group consisting of halogen, cyano, hydroxy, (C ₁ -C ₆)alkyl, (C ₁ -C ₆)alkoxy, (C ₁ -C ₂)alkyl substituted with one or more halogens, (C ₁ -C ₆)alkoxy substituted with one or more halogens, SR ^a , and NR ^b R ^c , in which R ^a and R ^c are as defined above.	
10	q. "heterocycle" or "heterocyclic ring" refers to any 3- or 4-membered ring containing a heteroatom selected from oxygen, nitrogen and sulfur; or a 5-, 6-, 7-, 8-, 9-, or 10- membered ring containing 1, 2, or 3 nitrogen atoms; 1 oxygen atom; 1 sulfur atom; 1 nitrogen and 1 sulfur atom; 1 nitrogen and 1 oxygen atom; 2 oxygen atoms in non-adjacent positions; 1 oxygen and 1 sulfur atom in non-adjacent positions; or 2 sulfur atoms in non-adjacent positions. The 5-membered ring has 0 to 1 double bonds, the 6- and 7-membered rings have 0 to 2 double bonds, and the 8, 9, or 10 membered rings may have 0, 1, 2, or 3 double bonds. The term "heterocyclic" also includes bicyclic groups in which any of the above heterocyclic rings is fused to a benzene ring, a cyclohexane or cyclopentane ring or another heterocyclic ring (for example, indolyl, quinolyl, isoquinolyl, tetrahydroquinolyl, benzofuranyl, dihydrobenzofuranyl or benzothienyl and the like). Heterocycles include: pyrrolidinyl, tetrahydrofuranyl, tetrahydrothiophenyl, piperidinyl, piperazinyl, azepane, azocane, morpholinyl, isochroman and quinolyl.	
15	r. "heterocyclic, optionally substituted" refers to a heterocyclic moiety as defined immediately above, in which up to 4 carbon atoms of the heterocyclic moiety may be substituted with a substituent, each substituent is independently selected from the group consisting of halogen, cyano, hydroxy, (C ₁ -C ₆)alkyl, (C ₁ -C ₆)alkoxy, (C ₁ -C ₂)alkyl substituted with one or more halogens, (C ₁ -C ₆)alkoxy substituted with one or more halogens, SR ^a , and NR ^b R ^c , in which R ^a and R ^c are as defined above. Any nitrogen atom within such a heterocyclic ring	
20		
25		
30		
35		

În acest exemplu, se regăsesc următoarele probleme (paragraful 35):

- Caracterele la subscript și superscript sunt prea mici.
- Caracterele la subscript sunt puse prea jos față de baza rândurilor.
- Caracterele la superscript sunt puse prea sus față de baza rândurilor.

În consecință, rândurile 34 și 35 sunt recunoscute astfel de procedura de OCR-are:

"Substituted with one or more halogens, (C -C) alkoxy substituted

1 2

8 8 9 8 9

with one or more halogens, SR, and NR R, in which R and R are"

Exemplul 15: o pagină ce conține caractere neobișnuite

	WO 2006/057705	PCT/
	c = speed of sound in water;	
	$\tilde{z}_u$ = initial altitude for beam pair u ;	
	$\Delta\epsilon_{z,p} = \epsilon_{z,p+1,u} - \epsilon_{z,p,u}$ = comparable to sway-reduced altitude difference;	
	$\Delta\epsilon_{\gamma,u} = \epsilon_{\gamma,p+1,u} - \epsilon_{\gamma,p,u}$ = comparable to sway-reduced horizontal displacem	
5	$\epsilon_{z,p,u}$ = difference of vertical linearization point in ping p , beam pair u , from nominal $\tilde{z}_u$;	
	$\epsilon_{z,p+1,u}$ = difference of vertical linearization point in ping $p+1$, beam pair u , nominal $\tilde{z}_u$;	
	$\epsilon_{\gamma,p,u}$ = difference of horizontal-range sample v linearization point in ping p	
10	u , from the nominal $\gamma_{v,u}$. Note that this is the same for all horizontal samples;	
	$\epsilon_{\gamma,p+1,u}$ = difference of horizontal-range sample v linearization point in ping pair u , from the nominal $\gamma_{v,u}$. Note that this is the same for all horizontal samples;	
15	$\gamma_{v,u}$ = nominal horizontal offset to horizontal-range sample u for beam pair	

În acest exemplu se regăsesc următoarele probleme:

- Caractere neobișnuite: grecești în italic, și chiar și caractere cu tildă.
- Subscriptul este și aici prea mic.

În cazul celor mai multe programe de OCR-are, toate caracterele neobișnuite nu vor fi recunoscute corect.

Anexa 2, pagina 18

Exemplul 16: un text cu fonturi și spațiere înguste

WO 2006/036330

PCT/US2005/028798

23. The method of claim 18, wherein the data is encoded onto the representative transmission symbol by using a modulation method selected from a group consisting of: amplitude modulation, phase modulation, frequency modulation, single-sideband modulation, vestigial-sideband modulation, quadrature amplitude modulation, orthogonal frequency division modulation, pulse-code modulation, pulse-width modulation, pulse-amplitude modulation, pulse-position modulation, pulse-density modulation, frequency-shift keying, and phase-shift keying.
24. The method of claim 18, wherein each of the at least two communication signals is transmitted through a communication medium selected from a group consisting of: a wire medium, a wireless medium, an optical fiber ribbon, a fiber optic cable, a single mode fiber optic cable, a multi-mode fiber optic cable, a twisted pair wire, an unshielded twisted pair wire, a plenum wire, a PVC wire, and a coaxial cable.
25. The method of claim 18, wherein the at least two communication signals are both transmitted wirelessly.
26. The method of claim 18, wherein the at least two communication signals are both transmitted through a wire medium.
27. The method of claim 18, wherein the at least two communication signals are transmitted through a wire medium, and wirelessly.

Acest exemplu nu este conform cu paragrafele 37 și 38. În consecință, programul de OCR-are nu poate să perceapă corect limitele cuvintelor, având drept rezultat faptul că procesul de OCR-are nu poate fi utilizat în acest caz.

Anexa 2, pagina 19

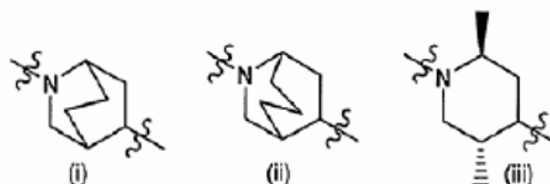
Exemplul 17: șampilă greșit aplicată de către oficiul de primire, înainte de scanare

WO 2006/058294

PCT/US2005/042931

reagent such as diborane or alkylation of the piperidine nitrogen with an alkyl halide or sulfonate ester provides the desired compounds.

[00176] Additionally, compounds of formulae (I, Ia, and Ib) in which the piperidine ring is replaced by:



Acest exemplu nu este conform cu paragraful 47. În consecință, primele șase cuvinte din textul de pe pagină nu pot fi citite cu programul OCR. Mai mult, șampilă introduce în plus caractere nejustificate, care vor polua motoarele de indexare în cazul în care pagina nu este verificată înainte de un operator, în privința calității.

Exemplul 18: o altă pagină cu formule matematice așezate greșit

WO 2006/079181 PCT/AU2006/000108

24

probability of the statistical outlier event of a noise only FFT bin magnitude being larger than a FFT bin containing both signal and noise is negligible.

Define,

$$\alpha = \sum_{n=0}^{N-1} r[n] \exp[-j2\pi(\frac{\hat{f}}{\hat{f}_s} - \frac{1}{2N})n] \quad (9)$$

$$\beta = \sum_{n=0}^{N-1} r[n] \exp[-j2\pi(\frac{\hat{f}}{\hat{f}_s} + \frac{1}{2N})n] \quad (10)$$

Then the discriminant, or distance metric, of frequency estimation error is defined as,

$$D(\varepsilon, \hat{\varepsilon}) = \frac{|\beta| - |\alpha|}{|\beta| + |\alpha|} \quad (11)$$

where, $\varepsilon = fT_s - \frac{k_{\max}}{N}$ (12)

and,

$$\hat{\varepsilon} = \hat{f}T_s - \frac{k_{\max}}{N}$$

For the initial frequency estimate using the FFT, $\hat{f}_0 T_s = \frac{k_{\max}}{N}$ and $\hat{\varepsilon} = 0$.

In the noiseless case,

$$D(\varepsilon, \hat{\varepsilon}) = \begin{cases} -1, & \varepsilon - \hat{\varepsilon} = -\frac{1}{2N} \\ 0, & \varepsilon - \hat{\varepsilon} = 0, \\ 1, & \varepsilon - \hat{\varepsilon} = \frac{1}{2N} \end{cases} \quad (13)$$

$D(\varepsilon, \hat{\varepsilon})$ is a monotonically increasing function of $\varepsilon - \hat{\varepsilon}$. Therefore, each $D(\varepsilon, \hat{\varepsilon})$, there is a unique inverse mapping to $\varepsilon - \hat{\varepsilon}$. Clearly, $D(\varepsilon, \hat{\varepsilon})$ may be used as a discriminant for fine frequency interpolation between FFT bin center frequencies.

There exists some functional relationship such that,

$$\hat{f}_1 T_s = \frac{k_{\max}}{N} + \psi[D(\varepsilon, \hat{\varepsilon})] \quad (14)$$

where, $\psi(\cdot)$ is a monotone increasing function. $\psi(\cdot)$ is called the frequency interpolation function and $\hat{f}_1$ is the first interpolated frequency estimate.

The requirement that $\hat{f}_1$ has zero error in the noiseless case is,

$$\psi[D(\varepsilon, \hat{\varepsilon})] = \varepsilon - \hat{\varepsilon}, \text{ for } -1 \leq D \leq 1. \text{ Therefore, } \psi^{-1}(\varepsilon - \hat{\varepsilon}) = D(\varepsilon, \hat{\varepsilon}).$$

THE FREQUENCY INTERPOLATION FUNCTION

Având în vedere că această pagină nu este conformă mai multor recomandări, rezultatul OCR-ării este inutilizabil:

- formule matematice încorporate, care nu sunt separate de paragrafele de text (paragraful 23);
- caractere neobișnuite în paragrafele de text (paragraful 40); stilul italic combinat cu caractere grecești (paragraful 39).

Modalitatea recomandată de a aranja această pagină este introducerea mai multor spații, pentru a separa formulele încorporate de paragrafe. Literele grecești nu trebuie să fie puse în italic în formule și paragrafe. Atunci când este posibil, se va evita utilizarea caracterelor circumflexe (^) pentru a marca variabile în paragrafele de text; se poate folosi atributul superscript în locul acestora: "epsilon circumflex" ar putea fi reprezentat ca $\epsilon^\wedge$ sau $\epsilon^{\text{circumflex}}$.

Exemplul 19: o pagină cu caractere italice și subliniate

WO 2006/038001

PCT/GB2005/003827

- 132 -

2-(3-([5-Chloro-4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino)piperidin-1-yl)-*N*-methylacetamide (S Enantiomer)

LCMS 399/401 [M+H]⁺, RT 1.88 min.

EXAMPLE 320

5 3-([5-Chloro-4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino)-*N*-isopropylpiperidine-1-carboxamide (Enantiomer 1)

LCMS 413/415 [M+H]⁺, RT 3.20 min.

EXAMPLE 321

10 3-([5-Chloro-4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino)-*N*-isopropylpiperidine-1-carboxamide (Enantiomer 2)

LCMS 413/415 [M+H]⁺, RT 3.19 min.

EXAMPLE 322

15 2-{3-[(4-([5-Chloro-4-(1*H*-indol-3-yl)pyrimidin-2-yl]amino)piperidin-1-yl)carbonyl]pyrrolidin-1-yl}-*N*-methylacetamide (Racemate)

LCMS (pH 5.8) 496/498 [M+H]⁺, RT 2.79 min.

Aceasta este o problemă frecvent întâlnită în publicația PCT, privind OCR-area. Această pagină nu este în conformitate cu următoarele recomandări:

- Paragraful 41: textul nu trebuie să fie subliniat. Sublinierea este în mod special nerecomandată în cazul formulelor chimice (dicționarele nu sunt de folos în aceste cazuri). În special, aceasta cauzează probleme cu toate caracterele care se intersectează cu linia: 1) y p ... nu sunt recunoscute corect.
- Paragraful 39: stilul italic nu este recomandat. Se recomandă pe cât posibil să nu se schimbe stilul fontului în cadrul unui cuvânt (programele de OCR-are deseori presupun că toate caracterele unui cuvânt au același stil). În consecință, "1*H*" și "*N*-" nu sunt recunoscute corect.

Anexa 2, pagina 22

Exemplul 20: o pagină complet ilizibilă

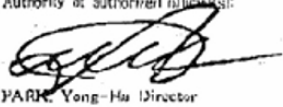
WO 2005/071074

PCT/KR2005/000214

HERZOGTUM TREATY OF THE INTERNATIONAL DEPOSITION OF THE DEPOSIT
OF MICROORGANISMS FOR THE PURPOSE OF PATENT PROCEDURE

**INTERNATIONAL FORM -
RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT**
issued pursuant to Rule 1.1

TO: Mokam Biotechnology Research Institute
4311, Huiung-ro, Kunsong-eup, Yongin-city, Kyonggi-do 449-910,
Republic of Korea

I. IDENTIFICATION OF THE MICROORGANISM	
Identification reference given by the DEPOSITOR: <i>Saccharomyces cerevisiae</i> HJ3501/MLK8 136	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: KCTC 10582BP
II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION	
The microorganism identified under I above was accompanied by: ☒ In a scientific description ☐ In a proposed taxonomic designation (Mark with a cross where applicable)	
III. RECEIPT AND ACCEPTANCE	
This International Depositary Authority accepts the microorganism identified under I above, which was received by it on January 13 2004 .	
IV. RECEIPT OF REQUEST FOR CONVERSION	
The microorganism identified under I above was received by this International Depositary Authority on _____ and a request to convert the original deposit to a deposit under the Budapest Treaty was received by it on _____.	
V. INTERNATIONAL DEPOSITARY AUTHORITY	
Name: Korean Collection for Type Cultures Address: Korea Research Institute of Bioscience and Biotechnology (KRIBB) #52, Oosong, Yusong-ku, Taejeon 305-380, Republic of Korea	Signature(s) of person(s) having the power to represent the International Depositary Authority of authorized official(s):  PARK Yong-Hu Director Date: January 17 2004

Please Print or Stamp Name RA

Această pagină nu trebuie acceptată de către oficii: a fost trimisă prin fax la 100 dpi, și nu este lizibilă nici măcar cu ochiul liber. Pentru a rezolva aceste cazuri, operatorii declară tot conținutul paginii ca fiind o imagine, având în vedere că nu poate fi extras textul.

[Sfârșitul Anexei 2 și al Standardului]